

STIC-Biotech/ChemLib

From: Saoud, Christine
Sent: Monday, May 06, 2002 10:30 AM
To: STIC-Biotech/ChemLib
Subject: sequence search

10/082902

Please search SEQ ID NO:1 and 2 in the patent and commercial databases. Please include interference search (pending).

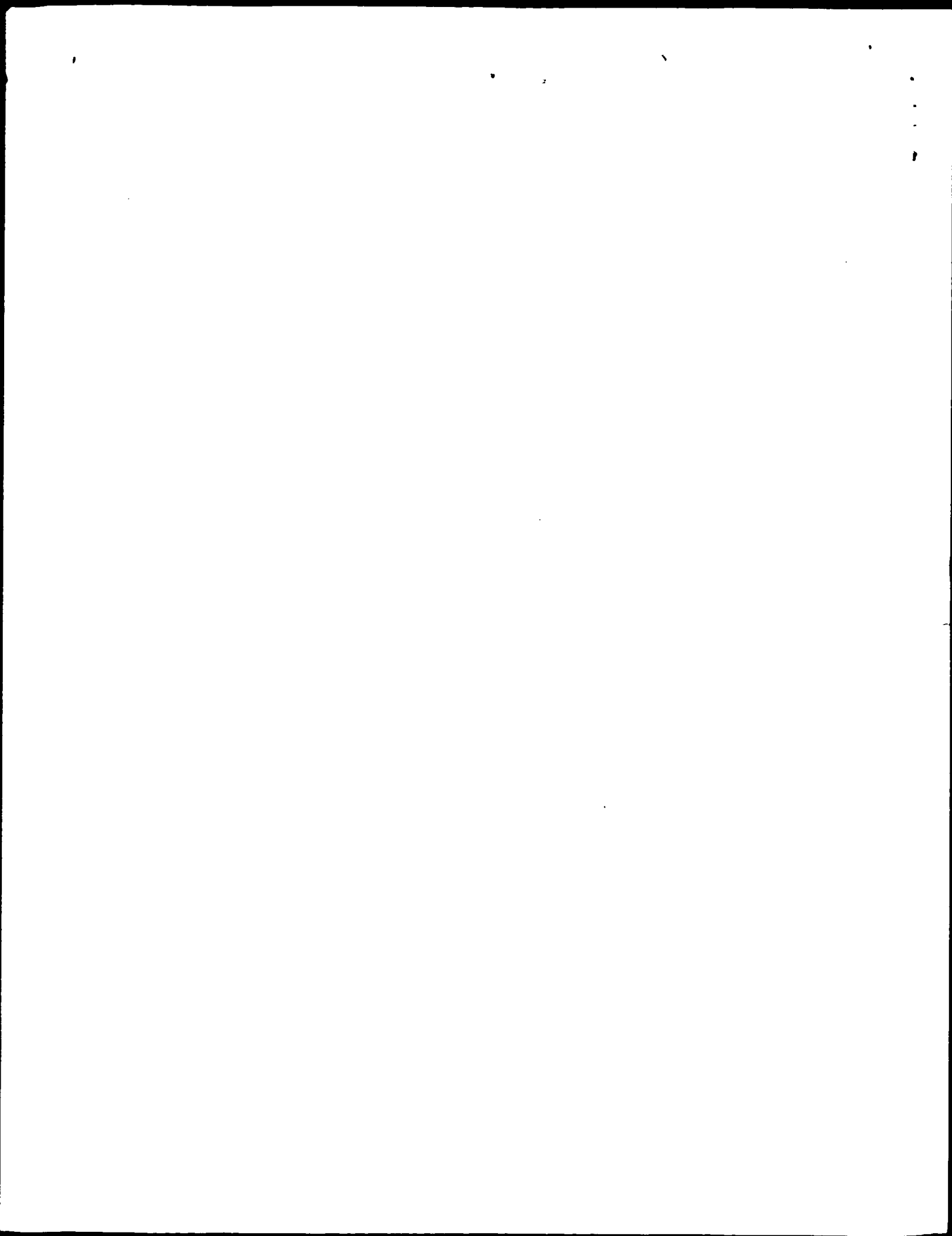
Thanks,
Christine Saoud
A.U. 1647
CM1 - 10E03
305-7519
mailbox 10C01

RECEIVED
MAY - 6 2002
(STIC)

if Contact: Sheppard
Searcher: _____
Phone: _____
Location: tel: 308-4499
Date Picked Up: _____
Date Completed: 5/11/02
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST(where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2002, 01:06:40 ; Search time 52.66 Seconds

(without alignments)
7752,428 Million cell updates/sec

Title: US-10-082-902-1

Perfect score: 1662

Sequence: 1 agtaagcagactgacgtac.....tgcttgggaaggtgtaga 1662

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*

2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*

3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*

4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*

5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*

6: /cgn2_6/ptodata/2/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	308.6	18.6	456	US-08-965-903B-15	Sequence 15, Appl
2	180.2	10.8	2135	US-08-965-903B-7	Sequence 7, Appl
3	173	10.4	1604	US-08-965-903B-5	Sequence 5, Appl
4	112.8	6.8	252	US-08-965-903B-12	Sequence 12, Appl
5	98.8	5.9	3254	US-08-965-903B-1	Sequence 1, Appl
6	97.2	5.8	372	US-08-965-903B-3	Sequence 3, Appl
7	64.8	3.9	300	US-08-965-903B-10	Sequence 10, Appl
8	52	3.1	7218	US-08-232-463-14	Sequence 14, Appl
9	51.4	3.1	3468	US-07-951-715A-2	Sequence 2, Appl
10	51.4	3.1	3468	US-08-459-448A-2	Sequence 2, Appl
11	51.4	3.1	3468	US-08-459-595A-2	Sequence 2, Appl
12	51.4	3.1	3468	US-08-459-504B-2	Sequence 2, Appl
13	51.4	3.1	3468	US-08-459-444-2	Sequence 2, Appl
14	51.4	3.1	3468	US-09-053-549-3	Sequence 3, Appl
15	51.4	3.1	3468	US-09-547-422-2	Sequence 2, Appl
16	44.4	2.7	1221	5212296-16	Sequence 2, Appl
17	44.4	2.7	1879	5212296-5	Sequence 2, Appl
18	41	2.5	846	US-07-901-707-12	Sequence 12, Appl
19	41	2.5	846	US-07-988-430-12	Sequence 12, Appl
20	41	2.5	846	US-08-425-336-12	Sequence 12, Appl
21	41	2.5	846	US-08-488-113B-12	Sequence 12, Appl
22	41	2.5	846	US-08-477-484B-12	Sequence 12, Appl
23	41	2.5	846	US-08-646-360-12	Sequence 12, Appl
24	41	2.5	846	US-08-839-765-12	Sequence 12, Appl
25	41	2.5	846	US-09-136-389-12	Sequence 12, Appl
26	41	2.5	846	PCT-US92-09487-12	Sequence 12, Appl
27	40	2.4	800	PCT-US95-04801-4	Sequence 4, Appl

ALIGNMENTS

C 28	39.8	2.4	2580	3	US-09-050-863-2	Sequence 2, Appl
C 29	39.8	2.4	2580	4	US-09-359-081-2	Sequence 2, Appl
C 30	39.8	2.4	4542	2	US-09-130-114-1	Sequence 1, Appl
C 31	39.8	2.4	9600	4	US-08-910-647-1	Sequence 1, Appl
C 32	39.8	2.4	10596	1	US-07-884-811-15	Sequence 15, Appl
C 33	39.8	2.4	10596	1	US-07-885-971-15	Sequence 15, Appl
C 34	39.8	2.4	10596	1	US-08-087-783A-15	Sequence 15, Appl
C 35	39.8	2.4	10596	1	US-08-194-088B-15	Sequence 15, Appl
C 36	39.8	2.4	10596	2	US-08-194-087-15	Sequence 15, Appl
C 37	39.8	2.4	10596	1	PCT-US93-04648-15	Sequence 15, Appl
C 38	39.4	2.4	1032	1	US-08-375-186-1	Sequence 1, Appl
C 39	39.4	2.4	1032	1	US-08-457-797A-4	Sequence 4, Appl
C 40	39.4	2.4	1032	1	US-08-812-025-4	Sequence 4, Appl
C 41	39.4	2.4	1032	4	US-09-138-873A-4	Sequence 4, Appl
C 42	39.4	2.4	1032	4	US-08-775-362-1	Sequence 1, Appl
C 43	39.2	2.4	4154	1	US-08-131-365B-37	Sequence 37, Appl
C 44	39.2	2.4	4154	2	US-08-668-123-37	Sequence 37, Appl
C 45	38.4	2.3	9551	1	US-08-056-200-93	Sequence 93, Appl

RESULT 1
US-08-965-903B-15
Sequence 15, Application US-08965903B

Patent No. 6060275

GENERAL INFORMATION:

APPLICANT: Hachem, Nit

TITLE OF INVENTION: PROTEIN AND CODING

SEQUENCE

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates

STREET: 350 Cambridge Ave., Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US-08-965,903B

FILING DATE: 07-NOV-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/030232

FILING DATE: 07-NOV-1996

ATTORNEY/AGENT INFORMATION:

NAME: Pellthory, Joanne R

REGISTRATION NUMBER: 42,995

REFERENCE/DOCKET NUMBER: 8600-0177.30

TELEPHONE: 650-324-0880

TELEFAX: 650-324-0960

TELEX:

INFORMATION FOR SEQ ID NO. 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 456 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: m-spry4

FEATURE:

NAME/KEY: Coding Sequence

LOCATION: 3..353

OTHER INFORMATION:

FILING DATE: 07-NOV-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/030232
 FILING DATE: 07-NOV-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Pelichory, Joanne R
 REGISTRATION NUMBER: 42,995
 REFERENCE/DOCKET NUMBER: 8600-0177.30
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-324-0880
 TELEFAX: 650-324-0960
 TELEX:
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1604 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: h-spry1 cDNA
 FEATURE:
 NAME/KEY: Coding Sequence
 LOCATION: 1..417
 OTHER INFORMATION:
 US-08-965-903B-5

Query Match 10.4%; Score 173; DB 3; Length 1604;
 Best Local Similarity 65.3%; Pred. No. 9.6e-35;
 Matches 254; Conservative 0; Mismatches 135; Indels 0; Gaps 0;
 QY 685 tgcgaagcctctgaggaagtgtaaatgcaagagatgcatcccccggagcgttcc 744
 DB 1 TGTGAACAGTGTGGAAAGTGCAGAGTGGAGATGACATGCTCCAGACCTACCATCC 60
 QY 745 tgcgtggtctgcaacagagatgctgtgtctcagccagactctgtgtaactatgacag 804
 DB 61 TGTGTGGCTGTAAACCGGACGAGTCCCTTGTGCTGCTGAGAGCATGGTGAATATGGAAC 120
 QY 805 tgcattgttctgaggaagcattctaccatgcaagaatgagagacatgaggtctcc 864
 DB 121 TGCATGTGTCTTAAGGACCATCTTCTACCACTGCTCCAAATGACGACGAGGATTC 180
 QY 865 tgcgtgcaacacccctgctccctgctccctcactgctgctgctgctgctccatc 924
 DB 181 TATTTCAGATATCTTCTGCTCTCTTACCAATACACTGCTGCTCTGATACCTGTATG 240
 QY 925 ggtgctctctcgtggtgtgtcctcctgctgctgctgctgctgctgctgctgctg 984
 DB 241 GGAGCATGCTTATTATTATTACTTCTGCTTACTGCTTATCTCTCTTAAGGATCC 300
 QY 985 aactgagccagagtgctgcaagcgtctgctgctgctgctgctgctgctgctgctg 1044
 DB 301 AACCTGTGAGAGGAGTGTATGACTGATCCCAATCCAGCCGAGGATGAGATGAACCTCC 360
 QY 1045 aacagcgtcactgcaagagcagcagcg 1073
 DB 361 AACACTGTCTATTGTAAAGCTGGAGAGCTG 389

RESULT 4

US-08-965-903B-12
 Sequence 12, Application US/08965903B
 Patent No. 6060275
 GENERAL INFORMATION:
 APPLICANT: Hachosen, Nlr
 APPLICANT: Krasnow, Mark A.
 TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates
 STREET: 350 Cambridge Ave., Suite 250
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/965,903B
 FILING DATE: 07-NOV-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/030232
 FILING DATE: 07-NOV-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Pelichory, Joanne R
 REGISTRATION NUMBER: 42,995
 REFERENCE/DOCKET NUMBER: 8600-0177.30
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-324-0880
 TELEFAX: 650-324-0960
 TELEX:
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 252 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: m-spry1
 FEATURE:
 NAME/KEY: Coding Sequence
 LOCATION: 2..250
 OTHER INFORMATION:
 US-08-965-903B-12

Query Match 6.8%; Score 112.8; DB 3; Length 252;
 Best Local Similarity 68.4%; Pred. No. 8.9e-20;
 Matches 156; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
 QY 669 caagacattctgctggtgagagcctgtggaagttaattgcaagagtgatgcatcc 728
 DB 25 CCAGCAAGATTCTATCTGTGGAAAGTGTGGCAAGTGAATGTGAGAGTATGACGCCCC 84
 QY 729 ccgagcgttctctctcctgctgctgctgcaacagagtgctgtgctcagccagactct 788
 DB 85 CCGGCGCATGCATCTCTGCTGCTGCTGATGCGGAGTGCCTCTCCGCGGAGAGCAT 144
 QY 789 gttcaactatgcaagtgcatgtgtgtgtgctgagggcattctaccactgcaagaatga 848
 DB 145 GGTGGAATACGGGACCTGATGCTGCTGCTCAAGGAGCATTTTCTACCAAGCTCCAAATGA 204
 QY 849 gtagcatgagggctctctgctgctgcaacacccctgctcctcctcctc 896
 DB 205 TGTGATGAGAGTCTTACTCGATNACCATGCTCTGTTACAGCT 252

RESULT 5

US-08-965-903B-1
 Sequence 1, Application US/08965903B
 Patent No. 6060275
 GENERAL INFORMATION:
 APPLICANT: Hachosen, Nlr
 APPLICANT: Krasnow, Mark A.
 TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
 NUMBER OF SEQUENCES: 20

QY 1547 gctaactgagtcaccagggtaccccttaagaagacccttgagagtcctctactactctctc 1606
Db 1174 yyy 1233
QY 1607 ccttcaccatcaccaccaccaccacttgccctgagtcctc 1648
Db 1234 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1275

RESULT 9

US-07-951-715A-2
Sequence 2, Application US/07951715A
Patent No. 5625136
GENERAL INFORMATION:
APPLICANT: Kozziel, Michael G.
APPLICANT: Desai, Nalin M.
APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.
APPLICANT: Warren, Gregory W.
APPLICANT: Evola, Stephen V.
APPLICANT: Crossland, Lyle D.
APPLICANT: Wright, Martha S.
APPLICANT: Merlin, Ellis J.
APPLICANT: Launis, Karen L.
APPLICANT: Rothstein, Steven J.
APPLICANT: Bowman, Cindy G.
APPLICANT: Dawson, John L.
APPLICANT: Dunder, Erik M.
APPLICANT: Pace, Gary M.
APPLICANT: Suttle, Janet L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.308
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/951,715A
FILING DATE: 25-SEP-1992
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Spruill, W. Murray
REGISTRATION NUMBER: 32,943
REFERENCE/DOCKET NUMBER: S-18805/A/C/C 1577/CIP
TELEPHONE: (919)541-8615
TELEFAX: (919)541-8615
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3468 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..3468
OTHER INFORMATION: /product= "Full-length pure maize
OTHER INFORMATION: optimized synthetic Bt"

OTHER INFORMATION: /note= "Disclosed in Figure 3 as syn1.mze"
US-07-951-715A-2

Query Match 3.1%; Score 51.4; DB 1; Length 3468;
Best Local Similarly 46.8%; Pred. No. 0.0011;
Matches 196; Conservative 0; Mismatches 221; Indels 2; Gaps 1;

QY 107 cgagcttcaggtttacacagacgtagggagatgctgtgacccctgacagctcctcaagac 166
Db 903 CAGCTCACACCTTACACACGACGCGGAGAGTACTGTGAGCGGCGCACAGAT 962
QY 167 cccctagaagccgtttcccgtagagtcagga--cctcagaccatagagcccccagat 224
Db 963 CATGGCAGACCCCGTGGGCTTACAGGCGCCCGAGTTCACCTTCCCTGTACGCGACCAT 1022
QY 225 cccacagagagcccccctgactcccaactcagatcagtcagccctcttgcacg 284
Db 1023 GGGCAGCGCCGCCCGCCAGCAGCCATGTCGCCAGCTGGGCGAGGCGGTACCGCAC 1082
QY 285 gatgcccagagccggtccagacccactacatctccttgcacagtggaagac 344
Db 1083 CCTGAGCAGCACCTGTACCGCCCGCTTCAACATGCGATCAACAACGACGTGAG 1142
QY 345 cagccatgtgagatgactacatagacaaacctagcctgagccctgacacagcccaaa 404
Db 1143 CGTGCTGACGACGACCGACCGGCTTGGCTTACGCGACGACGACGACCTGCCACGCGGTGA 1202
QY 405 gcgagaccggggcgggggccacagagctggcccgagcccgccgctgtgacagagatg 464
Db 1203 CCGCAGACGCGCGACCGGTGACGACCTGTGAGAGATCCCGCCGAGAACAAACGTGCC 1262
QY 465 caccacccatggtatcccttcagcggggcccccagctcgtgtgagcagcagcagca 523
Db 1263 CCCCCCGAGGCTTACACCGCCCTGTAGCCAGCATGTCCTCCGACGCGGCTTCA 1321

RESULT 10

US-08-459-448A-2
Sequence 2, Application US/08459448A
Patent No. 5859336
GENERAL INFORMATION:
APPLICANT: Kozziel, Michael G.
APPLICANT: Desai, Nalin M.
APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.
APPLICANT: Warren, Gregory W.
APPLICANT: Evola, Stephen V.
APPLICANT: Crossland, Lyle D.
APPLICANT: Wright, Martha S.
APPLICANT: Merlin, Ellis J.
APPLICANT: Launis, Karen L.
APPLICANT: Rothstein, Steven J.
APPLICANT: Bowman, Cindy G.
APPLICANT: Dawson, John L.
APPLICANT: Dunder, Erik M.
APPLICANT: Pace, Gary M.
APPLICANT: Suttle, Janet L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5859336artis Corporation
STREET: Patent & Trademark Dept., 520 White Plains
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS


```

Db 903 CAGCATACACATCTACACAGCCAGCCAGCCGAGTACTACTGAGGCGCCACAGAT 962
Qy 167 ccttagaagcgtttcttcctgaagtcagga--cctcagcccatggaagcccgat 224
Db 963 CATGCCACCCCGTGGGCTTACGGGGCCCGAGTTACCTTCCCTTACGGGACCAT 1022
Qy 225 ccacagagagccccccttactcccaactcagtcagtgtccagccctcttgaagccg 284
Db 1023 GGGCAAGCGCGCCGCCACAGAGCATGCTGGCCAGCTGGGGCCAGGGGCTTACCGCC 1082
Qy 285 gatgtccacagcgggtctccagcaaccactacacatcctaccatgagcaggtgaagc 344
Db 1083 CCGTACACACACCTCTTACCGCGCCCTTCAACATCGGCATCAACACACACAGCTGAG 1142
Qy 345 cagccatgtggaagatatacatagacaacaccttagccctgagccagccagcccaaa 404
Db 1143 CCGTGTGAGCGGACCGAGTGTGCTTACCGCACACAGCAACAGCTGCGCGCTGTA 1202
Qy 405 gcgagaccggggcgggcccgagagctggcccgagcccgccgctgtgacaggaagt 464
Db 1203 CGCGAAGAGCGGACCGTGTGAGAGAGCTGAGAGAGATCCCCCCAGAACACAGCTGCC 1262
Qy 465 caccacacatgtgattctccttagcgggcgcccgagctctgtgagcagcagagaaga 523
Db 1263 CCCCCCGCAGGGGCTTACGACCGCCGCTGAGCACGTCGATGATGTTCCGACGCGCTTCA 1321

```

```

RESULT 12
US-08-459-504B-2
Sequence 2, Application US/08459504B
Patent No. 6075185

```

```

GENERAL INFORMATION:
APPLICANT: Koziel, Michael G.
APPLICANT: Desai, Nalini M.
APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.
APPLICANT: Warren, Gregory W.
APPLICANT: Evola, Stephen V.
APPLICANT: Crossland, Lyle D.
APPLICANT: Wright, Martha S.
APPLICANT: Merrill, Ellis J.
APPLICANT: Launis, Karen L.
APPLICANT: Rothstein, Steven J.
APPLICANT: Bowman, Cindy G.
APPLICANT: Dawson, John L.
APPLICANT: Dunder, Erik M.
APPLICANT: Pace, Gary M.
APPLICANT: Suttle, Janet L.

```

```

TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6075185artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,504B
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/459,595
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992

```

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3468 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..3468
OTHER INFORMATION: /product= "Full-length pure maize
OTHER INFORMATION: optimized synthetic Bt."
OTHER INFORMATION: /note= "Disclosed in Figure 3 as syn1T.mze"
US-08-459-504B-2

```

```

Query Match 3.1%; Score 51.4; DB 3; Length 3468;
Best Local Similarity 46.8%; Pred. No. 0.0011;
Matches 196; Conservative 0; Mismatches 221; Indels 2; Gaps 1;

```

```

Qy 107 cggctcagagatttaccagagcgttgaggagatgtctgtgacccctgacagcctcctcaagc 166
Db 903 CAGCATACACATCTACACAGCCAGCCAGCCGAGTACTACTGAGGCGCCACAGAT 962
Qy 167 ccttagaagcgtttcttcctgaagtcagga--cctcagcccatggaagcccgat 224
Db 963 CATGCCACCCCGTGGGCTTACGGGGCCCGAGTTACCTTCCCTTACGGGACCAT 1022
Qy 225 ccacagagagccccccttgaactcccaactagtcagtgtccagccctcttgaagcgc 284
Db 1023 GGGCAAGCGCGCCGCCACAGAGCATGTGCGCCAGCTGGGCCAGCGCGCTGTA 1082
Qy 285 gatgtccacagcgggtcccgacacacacacacacacacacacacacacacacacacac 344
Db 1083 CCGTGTGAGCGGACCGTGTGAGAGAGCTGAGAGAGATCCCCCCAGAACACAGCTGCC 1142
Qy 345 cagccatgtggaagatatacatagacaacaccttagcctgagccctgagcagcccaaa 404
Db 1143 CCGTGTGAGCGGACCGAGTGTGCTTACCGCACACAGCAACAGCTGCGCGCTGTA 1202
Qy 405 gcgagaccggggcgggcccgagagctggcccgagcccgccgctgtgacaggaagt 464
Db 1203 CGCGAAGAGCGGACCGTGTGAGAGAGCTGAGAGAGATCCCCCCAGAACACAGCTGCC 1262
Qy 465 caccacacatgtgattctccttagcgggcgcccgagctctgtgagcagcagagaaga 523
Db 1263 CCCCCCGCAGGGGCTTACGACCGCCGCTGAGCACGTCGATGATGTTCCGACGCGCTTCA 1321

```

```

RESULT 13
US-08-459-444-2
Sequence 2, Application US/08459444A
Patent No. 6121014

```

```

GENERAL INFORMATION:
APPLICANT: Koziel, Michael G.
APPLICANT: Desai, Nalini M.
APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.
APPLICANT: Warren, Gregory W.
APPLICANT: Evola, Stephen V.
APPLICANT: Crossland, Lyle D.

```


Db	963	CATGGCCAGCCCGTGGGCTTTAGCGGCCCGCGAGTTACACTTCCCGCTAGCGGCACAT	1022
QY	225	cccccaagagcgccccctctgacctctccaaactcaatcagtcagtgtccagccccctctctgacagcgg	284
Db	1023	GGGCAACGGCGCCCCCGACAGCAGGCGATCTGGGCCAGCGGGCCAGGGCGCTGTACCGCAC	1082
QY	285	gattgtccacaagccgggtctcagcaccactcaccactcaccactctaccatctgacagcagtgaaagac	344
Db	1083	CCTGAGCAGCACCTGTACCGCGCGCCCTTTCAACATCGGCATCAACACACACACACTGAG	1142
QY	345	caagccatgtgtgagaaatgactactactaagaacaacctagtcctgtgcctctgacacacagcgccaaa	404
Db	1143	CGTCTGTGAGGCGCACCGAGTTCCCTACGCGACACAGCAGCAACATGGCCAGCGCGGTGA	1202
QY	405	gcggaaccggygggggggccccagaacttgacccccgaagcccgccgactgtgacaaagatgtc	464
Db	1203	CCGCAAGAGCGGCACCGTTGACAGCCTGTGACAGCGATGCCCGCCCGCAACAAACAAAGTCC	1262
QY	465	caaccacacttgatctcctctctagcggggccccagcctctgtgagcaagcagcagca	523
Db	1263	CCCCCGCAGGCGCTTACGACACCGCGCTGAACCACTGACAGCATGTTCCGACGCGCTTTA	1321

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2002, 02:16:20 ; Search time 200.42 Seconds

(without alignments)
14237.660 Million cell updates/sec

Title: US-10-082-902-1

Perfect score: 1662

Sequence: 1 aggtacgagctgagctgac.....tgcttgggaagtgtaga 1662

Scoring table: IDENTITY_NUC
Gapop 10.0, Capext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Genesec_032802:*

- 1: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1980.DAT:*
- 2: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1981.DAT:*
- 3: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1982.DAT:*
- 4: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1983.DAT:*
- 5: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1984.DAT:*
- 6: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1985.DAT:*
- 7: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1986.DAT:*
- 8: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1987.DAT:*
- 9: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1988.DAT:*
- 10: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1989.DAT:*
- 11: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1990.DAT:*
- 12: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1991.DAT:*
- 13: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1992.DAT:*
- 14: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1993.DAT:*
- 15: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1994.DAT:*
- 16: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1995.DAT:*
- 17: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1996.DAT:*
- 18: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1997.DAT:*
- 19: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1998.DAT:*
- 20: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA1999.DAT:*
- 21: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA2000.DAT:*
- 22: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA2001A.DAT:*
- 23: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA2001B.DAT:*
- 24: /SIDSL/gcgdata/hold-genesec/genesecn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	307.6	18.5	357	21 AAC07730	Human secreted pro
2	222	13.4	369	22 AAS27967	Novel cDNA encodin
3	189	11.4	2138	21 AAA14073	Human SPROUTY-1 pa
4	189	11.4	2195	21 AAA14072	Human SPROUTY-1 pa
5	189	11.4	2330	22 AAF32646	Human cDNA encodin
6	189	11.4	2393	22 AA157844	Human polynucleoti
7	189	11.4	2555	21 AAC75393	Human ORFX ORF948
8	180.2	10.8	1508	21 AAA98830	Human ORFX ORF948
9	180.2	10.8	1508	21 AAA98111	Human differentiat

10	180.2	10.8	2135	19 AAV32399	Homo sapiens sply2
11	179.6	10.8	2199	22 AA159630	Human polynucleoti
12	179.4	10.8	1771	21 AAA14075	Human SPROUTY-1 pa
13	162.6	9.8	1603	19 AAV32400	Homo sapiens sply1
14	159.6	9.6	1622	22 ABA09662	Human bone marrow
15	156.8	9.4	723	22 ABA09706	Human bone marrow
16	114	6.9	582	22 ABA59853	Human foetal liver
17	114	6.9	582	22 ABA28317	Probe #6783 for ge
18	114	6.9	582	22 AAK08124	Human brain expres
19	114	6.9	582	22 AAK34002	Human bone marrow
20	114	6.9	582	22 AA139725	Probe #8411 used t
21	98.8	5.9	3254	19 AAV32402	Drosophila melanog
22	97.4	5.9	1057	21 AACT79934	Human secreted pro
23	97.2	5.8	2122	23 AB103707	Drosophila melanog
24	97.2	5.8	4122	23 AB103706	Drosophila melanog
25	96.8	5.8	286	22 ABA72394	Human foetal liver
26	96.8	5.8	286	22 ABA83196	Probe #16662 for g
27	96.8	5.8	286	22 AAK20818	Human brain expres
28	96.8	5.8	286	22 AAK46970	Human bone marrow
29	96.8	5.8	286	22 AA152806	Probe #21492 used
30	91	5.5	601	22 AAB81506	Human differentiat
31	79.4	4.8	760	21 AAA14071	Human SPROUTY-1 pa
32	79.4	4.8	2030	21 AAA14074	Human SPROUTY-1 pa
33	64.8	3.9	300	19 AAV32401	Homo sapiens sply3
34	64.4	3.9	255	22 AAB82143	Rat differentiat t
35	51.4	3.1	3474	14 AAO39018	Sequence of pure m
36	51.4	3.1	3474	18 AAT74012	Maize optimised-B.
37	51.4	3.0	3474	19 AAV16191	Maize optimised DN
38	50.2	3.0	795	19 AAV55830	Flga insert stablil
39	47.2	2.8	58857	21 AA558471	Nucleotide sequenc
40	46.2	2.8	493	21 AA250580	Human epidermal pr
41	45	2.7	1337	20 AAZ17263	Human gene express
42	44.6	2.7	1459	21 AAA02528	Human colon cancer
43	44.6	2.7	6652	23 ABL13170	Drosophila melanog
44	44.4	2.7	1879	12 AAO11126	Sequence encoding
45	44.2	2.7	10732	21 AAA10594	Gene encoding a su

ALIGNMENTS

RESULT 1	
ID AAC07730	standard; cDNA: 357 BP.
XX AAC07730;	
XX	
XX	
XX	06-OCT-2000 (first entry)
XX	
XX	Human secreted protein 5' EST, SEQ ID NO: 11805.
XX	
XX	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW	gene therapy; chromosome mapping; ss.
KW	
XX	
OS	Homo sapiens.
OS	
XX	
PN	EP1033401-A2.
XX	
PD	06-SEP-2000.
XX	
XX	21-FEB-2000; 2000EP-0200610.
XX	
XX	26-FEB-1999; 99US-0122487.
XX	
XX	(GEST) GENSET.
XX	
XX	Dumas Milne Edwards J, Duclet A, Giordano J;
XX	
XX	WPI; 2000-500381/45.
XX	
XX	New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX	obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX	diagnostic, forensic, gene therapy and chromosome mapping procedures -
PT	
PT	

XX Claim 1: SEQ ID 11805; 71bp + CD-ROM; English.
PS The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or poly(A)-RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 357 BP; 69 A; 124 C; 98 G; 65 T; 1 other;

Query Match 18.5%; Score 307.6; DB 21; Length 357;
Best Local Similarity 99.7%; Pred. No. 3.2e-67;
Matches 307; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 agttagcagagctgagctgacagcgagagctgagctgagagcgagagccttgccg 60
Db 50 agttagcagagctgagctgacagcgagagctgagctgagagcgagagccttgccg 109
QY 61 gtcttcacgagcgagctgagctgagagcgagagcgagagccttgagattt 120
Db 110 gtcttcacgagcgagctgagctgagagcgagagcgagagccttgagattt 169
QY 121 acacagacgttgggagctgagctgagagcgagagccttcctcaaaagcgagagcctgt 180
Db 170 acacagacgttgggagctgagctgagagcgagagccttcctcaaaagcgagagcctgt 229
QY 181 ttcttcgttacagctcagagcctcagagccttcagagccttcagagccttcag 240
Db 230 ttcttcgttacagctcagagcctcagagccttcagagccttcagagccttcag 289
QY 241 ttgagctcagagctgagctgagagccttcagagccttcagagccttcagagccttcag 300
Db 290 ttgagctcagagctgagctgagagccttcagagccttcagagccttcagagccttcag 349
QY 301 ctccagca 308
Db 350 ctccagca 357

RESULT 2
AAS27967/c
ID AAS27967 standard; cDNA; 369 BP.

XX AAS27967;

XX 07-NOV-2001 (first entry)

DE Novel cDNA encoding for human respiratory antigen #99.

XX Human: respiratory antigen; respiratory disorder; throat disorder;
KM lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
KM anti allergic; anti asthmatic; anti inflammatory; olfactory;
KW respiratory active; ss.

XX Homo sapiens.

XX WO200155448-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01333.

XX

PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 14-AUG-2000; 2000US-0225799.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226682.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0234984.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.

[illegible]

PS Claim 1a: Page 24-25; 84pp; German.

CC This invention describes novel human nucleic acid sequences (I) from
CC endothelial cells which have angiogenic, antipsoriatic, antidiabetic,
CC ophthalmological, nephrotropic, cytostatic, hepatotropic,
CC antiatherosclerotic, antiarthritic, antirheumatic activity and can be
CC used for gene therapy. The nucleic acid sequences may find use as tools
CC to identify agents against angiogenic diseases. The polypeptides may be
CC used in medicaments for gene therapy to treat angiogenic diseases. (I)
CC in sense or antisense form may be used. The genomic genes may also be
CC used with suitable promoters and/or enhancers. The sequences (nucleic
CC acid and proteins) may be used to treat psoriasis, arthritis, especially
CC rheumatoid arthritis, haemangioma, angiodioma, eye diseases, especially
CC diabetic retinopathy, neovascular glaucoma, kidney disease, such as
CC glomerulonephritis, diabetic nephropathy, malign nephrosclerosis,
CC thrombotic microangiopathic syndrome, transplantation rejection and
CC glomerulopathy, fibrotic disease, such as liver cirrhosis, mesangial cell
CC proliferative disease, arteriosclerosis and injury to nerve tissue.

XX Sequence 1508 BP; 429 A; 306 C; 312 G; 461 T; 0 other:

Query Match 10.8%; Score 180.2; DB 21; Length 1508;
Best Local Similarity 66.7%; Pred. No. 3.9e-35;
Matches 273; Conservative 0; Mismatches 133; Indels 3; Gaps 1;

QY 661 gagctggagaacactctctctgtgagagcctgtggaatgaaatgcaagagatgt 720
DB 280 gatttggagctgacagcctacagctgtgagactgtgcaaatgtaaagatgac 339
QY 721 gataccccccgagctgctccctcctgagctgcaacagagatgctgtgctcagcc 780
DB 340 acctaccagaagcctctgcatcagactgagatctgcaacagcagtgcttgcctgcgcg 399
QY- 781 cagactgtgtcaactatgagcagtgatgtgtgtgagagagcattctacacatgc 840
DB 400 cagaacgttatgactatgagactgtgtatgtgtgtaagagctctctctacatcgt 459
QY 841 agaatgagagagatgagagctcctcctgagcctccctgctcctcctcctcctcctc 900
DB 460 tctaat--gactgagagagacactgtgtcgaacaacctgtctcttgagccagctcac 516
QY 901 tgcctgcccgcgtgtcctcctcaatgggtgtgtctcctcgtgtgtcctcctgtcgtc 960
DB 517 tgtgtacacagatgtgcagcagtggtgtgtcctcctcctcctcctcctgttatagtgt 576
QY 961 tacctgtcgtccacgcgtgtgtgagctgtgcccagcgtgtgtaagacgtctgcgcgc 1020
DB 577 taccttcacagcagaaggtgtgtccttaattgtgcacaggggtgttatgacgggttaacag 636
QY 1021 cctgtgtcgcgtgtcagacagcagaacagcagctcattctgcaagaacgca 1069
DB 637 cctgtgtcgcgtgttataaactcaacacagtttctgcaagaatgccca 685

RESULT 9

AAA98111 standard; CDNA: 1508 BP.

AAA98111:

26-JAN-2001 (first entry)

Human differentiation-associated protein EST encoded CDNA #6.

XX Proliferation-associated protein; differentiation-associated protein;
XX EST; expressed sequence tag; endothelial cell; anti-angiogenic;
XX angiogenic; antipsoriatic; anti-arthritis; antifibrotic; hepatotropic;
XX antiatherosclerotic; nephrotropic; antidiabetic; ophthalmological;
XX immunosuppressive; neuroprotective; gene therapy; psoriasis; arthritis;
XX hemangioma; diabetic retinopathy; glomerulonephritis; liver cirrhosis;
XX transplant rejection; arteriosclerosis; nervous tissue injury; human; ss.

OS Homo sapiens.

XX W0200053734-A2.

XX 14-SEP-2000.

XX 08-MAR-2000; 2000WO-EP02005.

XX 09-MAR-1999; 99DE-1011684.

XX 01-OCT-1999; 99DE-1048679.

XX (SCHD) SCHERING AG.

XX Thierach K, Glienke J, Hinzmann B, Pilarsky C;

XX WPI: 2000-572267/53.

XX Nucleic acid sequences from human endothelial cells, useful for gene
XX therapy of angiogenesis and for identifying antiangiogenic agents -

XX Claim 1a: Page 53-54; 115pp; German.

CC This invention describes novel nucleic acid sequence (I) obtained from
CC an endothelial cell comprising any of 59 sequences (defined and given in
CC the specification), its allelic variant, or complement. The products of
CC the invention have anti-angiogenic, angiogenic, antipsoriatic,
CC anti-arthritis, antifibrotic, antiatherosclerotic, nephrotropic,
CC antidiabetic, ophthalmological, hepatotropic, immunosuppressive and
CC neuroprotective activity. (I) are implicated in differentiation of
CC proliferation of endothelial cells. (I) are used for expression of
CC polypeptides (IV) that are useful for identifying agents (A) for
CC treatment of angiogenic disease, to express (IV) from gene therapy
CC vectors for treatment of such diseases and as antisense reagents. (IV)
CC are also used to raise specific antibodies. Angiogenic diseases that
CC may be treated include psoriasis, arthritis, hemangioma, diabetic
CC retinopathy, glomerulonephritis, transplant rejection, liver cirrhosis,
CC arteriosclerosis and injuries to nervous tissue. AAA98102-898161
CC represent proliferation-associated and differentiation-associated EST
CC (expressed sequence tag) derived sequences which are described in the
CC method of the invention.

XX Sequence 1508 BP; 429 A; 306 C; 312 G; 461 T; 0 other;

Query Match 10.8%; Score 180.2; DB 21; Length 1508;
Best Local Similarity 66.7%; Pred. No. 3.9e-35;
Matches 273; Conservative 0; Mismatches 133; Indels 3; Gaps 1;

QY 661 gagctggagaacactctctctgtgagagcctgtggaatgaaatgcaagagatgt 720
DB 280 gatttggagctgacagcctacagctgtgagactgtgcaaatgtaaagatgac 339
QY 721 gataccccccgagctgctccctcctgagctgcaacagagatgctgtgctcagcc 780
DB 340 acctaccagaagcctctgcatcagactgagatctgcaacagcagtgcttgcctgcgcg 399
QY 781 cagactgtgtcaactatgagcagtgatgtgtgtgtgagagagcattctacacatgc 840
DB 400 cagaacgttatgactatgagactgtgtatgtgtgtgtaagagctctctctacatcgt 459
QY 841 agaatgagagagatgagagctcctcctgagcctccctgctcctcctcctcctcctc 900
DB 460 tctaat--gactgagagagacactgtgtcgaacaacctgtctcttgagccagctcac 516
QY 901 tgcctgcccgcgtgtcctcctcaatgggtgtgtctcctcgtgtgtcctcctgtcgtc 960
DB 517 tgtgtacacagatgtgcagcagtggtgtgtcctcctcctcctcctcctgttatagtgt 576
QY 961 tacctgtcgtccacgcgtgtgtgagctgtgcccagcgtgtgtaagacgtctgcgcgc 1020
DB 577 taccttcacagcagaaggtgtgtccttaattgtgcacaggggtgttatgacgggttaacag 636
QY 1021 cctgtgtcgcgtgtcagacagcagaacagcagctcattctgcaagaacgca 1069

DB 182 gagagcatggtggaataatgaaacctgcatgtctagtaagaaggcatctctaccactgac 241
 QY 841 acgaatgagacgacatgagggctcctgctgacacacccctgctcctgctccgctccac 900
 XX 1 1111 1111 11 1111 11 11 11111111 11 11
 DB 242 tccaatgagacgaggaaggtatctcattcctaataacacctgctggttcaacatcacac 301
 QY 901 tgcctgcccgcctggtctcctcctcctcctcctcctcctcctcctcctcctcctcctc 960
 XX 11111111 11 11111111 11 11 11 11111111 11111111
 DB 302 tgcctgctcctgacatctgctgtaaggacacgactcttatttctacctgcttactctgt 361
 QY 961 taccctgctcctgacacgctgctgtaagctgacacgctgctgctgctgctgctgctgctg 1020
 XX 11 111111 11 1111111111 11 11 111111 11 1111
 DB 362 tatctcctcctgctcctgacacgctgctgctgctgctgctgctgctgctgctgctgctg 421
 QY 1021 cctggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1073
 XX 11 1111 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 DB 422 ccagggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 474

RESULT 13
 AAV32400
 ID AAV32400 standard; cDNA to mRNA; 1603 BP.
 AC AAV32400;
 XX
 DT 11-SEP-1998 (first entry)
 XX
 DE Homo sapiens spry1 gene.
 XX
 KW spry1; sprouty 1 protein; angiogenesis; treatment; cancer;
 KW lateral inhibitory signal; epithelial tube branching; inhibition;
 KW neovascularisation; diabetic retinopathy; retrolental fibroplasia;
 KW neovascular glaucoma; psoriasis; angiodiomas; immune; non-immune;
 KW inflammation; rheumatoid arthritis; capillary proliferation;
 KW atherosclerotic plaques; haemangiomas; Kaposi's Sarcoma;
 KW endometriosis; scar formation; wound healing; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..420
 FT /tag= a
 FT /product= sprouty 1 protein
 FT
 XX
 PN W09820032-A1.
 XX
 PD 14-MAY-1998.
 XX
 PF 07-NOV-1997; 97WO-US20398.
 XX
 PR 07-NOV-1996; 96US-0030232.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 PI Hachon N, Krasnow MA;
 XX
 DR WPI: 1998-286862/25.
 DR P-PSDB: AAW48793.
 XX
 PT New isolated human sprouty polypeptide(s) - used to develop products
 PT for treating disorders involving angiogenic dysfunction, e.g.
 PT cancers, diabetic retinopathy, glaucoma, psoriasis or inflammation
 XX
 PS Claim 14; Page 50-51; 90pp; English.
 XX
 CC The sequence is that encoding human sprouty 1 protein. The protein
 CC functions as a lateral inhibitory signal during epithelial tube
 CC branching. It can be used for inhibiting angiogenesis and
 CC neovascularisation in the treatment of e.g. cancers, diabetic
 CC retinopathy, retrolental fibroplasia, neovascular glaucoma, psoriasis,
 CC angiodiomas, immune and non-immune inflammation (including rheumatoid
 CC arthritis), capillary proliferation within atherosclerotic plaques,
 CC haemangiomas, Kaposi's Sarcoma, endometriosis and unwanted scar

CC formation in wound healing. Products of the DNA and proteins
 CC can also be used for detection and diagnosis.
 XX
 SQ Sequence 1603 BP; 411 A; 320 C; 316 G; 556 T; 0 other;

Query Match 9.88; Score 162.6; DB 19; Length 1603;
 Best Local Similarity 65.38; Pred. No. 1e-30;
 Matches 254; Conservative 0; Mismatches 134; Indels 1; Gaps 1;

QY 685 tgcgagggcctgtgggaattgaatgaagagtgatgaccccccagctgctcc 744
 DB 1 tgtgacacgctgtgggaattgaatgagtgatgagacgactgtcccaagaccctaccac 60
 QY 745 tgcctgctcctgacacacgagtgctgctgctgctgctgctgctgctgctgctgctgctg 804
 DB 61 tgcctgctcctgacacacgagtgctgctgctgctgctgctgctgctgctgctgctgctg 120
 QY 805 tgcctgctcctgacacacgagtgctgctgctgctgctgctgctgctgctgctgctgctg 864
 DB 121 tgcctgctcctgacacacgagtgctgctgctgctgctgctgctgctgctgctgctgctg 180
 QY 865 tgcctgctcctgacacacgagtgctgctgctgctgctgctgctgctgctgctgctgctg 924
 DB 181 taccagataatcctcctgctcctgctcctgctcctgctcctgctcctgctcctgctcctg 240
 QY 925 ggtgctcctcctgctgctgctcctgctcctgctcctgctcctgctcctgctcctgctcctg 984
 DB 241 gagcagcatgcttatttctacactgcttactgcttactgcttactgcttactgcttactgct 300
 QY 985 aagctgcccacgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1044
 DB 301 aagct-gtcaggaggtgtgtatgactgactgactgactgactgactgactgactgactg 359
 QY 1045 aacagcgtcatctgcaagaagcagcagcagcagcagcagcagcagcagcagcagcagc 1073
 DB 360 aacactgtctattgtacagctgagagctg 388

RESULT 14
 ABA09622
 ID ABA09622 standard; DNA; 1622 BP.
 XX
 AC ABA09622;
 XX
 DT 15-JAN-2002 (first entry)
 XX
 DE Human bone marrow expressed oligonucleotide SEQ ID NO: 48.
 XX
 KW Human; bone marrow; cytostatic; antirheumatic; antiarthritic; vulnery;
 KW antinflammatory; antibacterial; immunosuppressive; vasotropic; cancer;
 KW antiparkinsonian; neuroprotective; nootropic; haemostatic; osteopathic;
 KW anticancer; fungicide; antidiabetic; antisthmatic; anti allergic;
 KW immunostimulant; analgesic; cerebroprotective; antianemic; infection;
 KW nervous system disorder; autoimmune disorder; inflammation; allergy; ds.
 XX
 OS Homo sapiens.
 XX
 PN W0200174836-A1.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US10472.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PR 23-AUG-2000; 2000US-0649267.
 PR 30-NOV-2000; 2000US-250583P.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Dermanac RT, Ford JE, Boyle BJ.


```

Db 329 atgatgaagacaactgtgtgtatgagccctgtcttgtgtgacctagtcttctgttgtcc 388
Qy 911 gctgtccttcattggtgtctctctccgtgtgtgtcctgtcctgtctctgtacctgtc 970
    ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||
Db 389 gctgggcaagccattgagcctcatctccctctctctaccctgtgtgtgtacctgtccta 448
Qy 971 caaccggctgtcgttgaagctgtgcccagcgtgtgtatgagccgtctgtgcccgtgtgtcc 1030
    || ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 449 cccgtgtgtgtccttcattctgtgtcccaagggctatgtatgacctccggtgacacagctgtcc 508
Qy 1031 gctgcaagcacacgaacagcgtcatctgtcaca 1062
    ||||| ||||| | ||| | |||||
Db 509 gctgcaagagagcacaccacactgtgtgtcaga 540

```

Search completed: May 11, 2002, 05:34:20
 Job time: 11880 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2002, 01:05:50 ; Search time 1753.28 Seconds
(without alignments)
19837.050 Million cell updates/sec

Title: us-10-082-902-1

Perfect score: 1662

Sequence: 1 aggaagcagagctgagctgac.....tgctctggggaaggtgtaga 1662

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBank:*

1: gb_ba:*

2: gb_hlg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
------------	-------------	--------	-------	-------------

1	1591.2	95.7	4934	9	AF227516	AF227516 Homo sapi
2	888	53.4	1405	10	AF176906	AF176906 Mus muscu
3	869.8	52.3	1131	10	AB019280	AB019280 Mus muscu
4	827.8	49.8	155614	2	AC012620	AC012620 Homo sapi
5	822.4	49.5	183042	2	AC099211	AC099211 Rattus no
6	734.2	44.2	7053	3	AF227517	AF227517 Homo sapi
7	291.8	17.6	2890	5	AF373368	AF373368 Danto rer
8	205	12.3	2502	10	AF176903	AF176903 Mus muscu
9	201.8	12.1	176250	2	AL627086	AL627086 Mus muscu
10	194.6	11.7	1889	5	AF176904	AF176904 Gallus ga
11	189	11.4	2330	6	AX058191	AX058191 Sequence
12	189	11.4	2520	9	AK026960	AK026960 Homo sapi
13	189	11.4	156451	2	AC025511	AC025511 Homo sapi
14	189	11.4	176959	9	AC026402	AC026402 Homo sapi
15	180.2	10.8	1508	6	AX035193	AX035193 Sequence
16	180.2	10.8	1508	6	AX342047	AX342047 Sequence
17	180.2	10.8	1508	6	AX342107	AX342107 Sequence
18	180.2	10.8	1918	9	BC015745	BC015745 Homo sapi
19	180.2	10.8	2135	9	AF039843	AF039843 Homo sapi
20	180.2	10.8	141175	2	AC012110	AC012110 Homo sapi
21	180.2	10.8	179470	2	AL162211	AL162211 Homo sapi
22	180.2	10.8	191652	9	AL154668	AL154668 Human DNA
23	173	10.4	1604	9	AF041037	AF041037 Homo sapi
24	169	10.2	948	10	AF153084	AF153084 Mus muscu
25	167.4	10.1	1519	10	AF176905	AF176905 Mus muscu
26	164	9.9	155614	2	AC012620	AC012620 Homo sapi
27	161.2	9.7	168066	9	AC025226	AC025226 Homo sapi
28	159.6	9.6	135484	2	AL156301	AL156301 Homo sapi
29	159.6	9.6	196456	2	AC009620	AC009620 Homo sapi
30	159.6	9.6	240000	9	HSAR271735	HSAR271735 Homo sapi
31	155.6	9.4	945	5	AF331824	AF331824 Xenopus l
32	155	9.3	924	5	AF331825	AF331825 Xenopus l
33	98.8	5.9	3254	3	AF039842	AF039842 Drosophi
34	98	5.9	73889	3	AC009983	AC009983 Drosophi
35	97.2	5.8	91318	2	AC015391	AC015391 Drosophi
36	97.2	5.8	113816	2	AC004358	AC004358 Drosophi
37	97.2	5.8	172271	3	AC010018	AC010018 Drosophi
38	97.2	5.8	178018	3	AC091206	AC091206 Drosophi
39	91	5.5	601	6	AX210373	AX210373 Sequence
40	83.8	5.0	1465	5	AF177875	AF177875 Gallus ga
41	69	4.2	125020	9	AF429315	AF429315 Homo sapi
42	64.8	3.9	300	9	AF041038	AF041038 Homo sapi
43	64.4	3.9	255	6	AX211010	AX211010 Sequence
44	63.6	3.8	125020	9	AF429315	AF429315 Homo sapi
45	62	3.7	46448	9	AC005789	AC005789 Homo sapi

ALIGNMENTS

RESULT	1
LOCUS	AF227516
DEFINITION	Homo sapiens sprouty-4A mRNA, complete cds.
ACCESSION	AF227516
VERSION	AF227516.1 GI:12655912
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE	1 (bases 1 to 4934)
AUTHORS	Leeksa,O.C., van Achterberg,T.A.E., Spaargaren,M., Kr. von dem Borne,A.E.G., Pannekoek,H. and de Vries,C.J.M.
TITLE	Identification of a novel human Sprouty homolog
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 4934)
AUTHORS	de Vries,C.J.M., Leeksa,O.C., van Achterberg,T.A.E. and Pannekoek,H.
TITLE	Direct Submission
JOURNAL	Submitted (21-JAN-2000) Biochemistry, Academic Medical Center, Meibergdreef 15, Amsterdam 1105 AZ, The Netherlands
FEATURES	Location/Qualifiers

```

source
1. .4934
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="smooth muscle"
/tissue_type="umbilical artery"
188..1156
/codon_start=1
/product="sprouty-4A"
/protein_id="AAK0652.1"
/db_xref="GI:1265913"
/translation="MSPLEPTGPLEACFSVQSRSSPMEPPIDQSAPELPNVSVMQPL
LSDMSHSRLQHPLEPLIDQVKTSHENDYIDNPISALTGKRTGGAPELAPPA
RCDQDYNHMSISGRSVSSSTSDRLDHPAPVADQASRAVRIDPKYVH
COPDLKPAVPELDHFLFLCACGSKCKECSAPPTLSCWVCWCECSAOTLVN
YGTMCILVQGIIFYHCTNEDEGSCADHPSCSSNCCARMSFNGALSYVLPCLLCTLP
ATGCVKLAQSGYRLRPRGCRKHTNSVICKAASGAKTSRPRPKPF"

BASE COUNT      1103 a 1365 c 1208 g 1258 t
ORIGIN

Query Match      95.7%; Score 1591.2; DB 9; Length 4934;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 1654; Conservative 0; Mismatches 3; Indels 51; Gaps 1;

QY 6 ggaagctgagctgacagcgagagctgagctgagagcgagagagcttgcggttc 65
Db 1 GCGAGCTGAGCTGACAGCGCGAGCTGCGCTGTGAGCGAGCGAGAGCTTGCGCGTTC 60

QY 66 tcgagcggcgctctgagagctagcggcggttaacctgcccggcttcagattaca 125
Db 61 TCCGACCGCGCTTGTGCGAGTACAGCGGGGCTACGCTGCCCCGCTTACGATTACACA 120

QY 126 gaagtgagagctgctgtgagcttgcagcttctcaaa----- 163
Db 121 GACGTGGGGCGATGCTGTGACCTCGACGCTCTCAACACGCTGTATTGAGCGCTTG 180

QY 164 -----ggcccttaagagctgttcttcgtaagtc 194
Db 181 CAGCTGATGCTACAGCCCCCTCCACAGGGGCCCTTAGAAGCCGTGTTCTCCGTACGTC 240

QY 195 caagacctcagacccttgagagcccgatcccaagagcgcccgcttgcagcttc 254
Db 241 CAGGACCTCCAGCGCCCATMGAGCGCCCGATGCCACAGAGCGCCCTTGACTCCCAACTC 300

QY 255 agtcatgtccagcccttcttgcagcgagtgctccacagcggttcacagaccact 314
Db 301 AGTATGCTCCAGCCCTTCTTGACAGCCGAGATGTCACACAGCGCGCTCCAGCACCCACT 360

QY 315 caacatctaaccttgcagaggtgagagacacacacatgtgagatgacataagacaa 374
Db 361 CACCATCTTACCATTTACACAGAGTGAAGACACGACATGTGAAGATGACTACATACACAA 420

QY 375 cccctagctgagcccttgacacccgacccaaagcgagacggcgagggcgcccaagctgc 434
Db 421 CCTTAGCTGGCCCTTGACACCGGGCCCAAGCGGACCGGGGGGGGGCCCCCAAGAGCTGCG 480

QY 435 ccgagagccggccgctgtgtaacagagtgcaaccacatgtgacttcttcagggcg 494
Db 481 CCGGACCGCGCCCGCTGTACACAGAGTGTACCCACCATTTGATCTCCCTTACGCGGCG 540

QY 495 ccccaagctctgtagagcagcagcagacacatctctgcagcaacggctcttaacacat 554
Db 541 CCCCAGCTGTGTAGAGCAGCAGCAGCAGCATCTCTTGACCAACGGCTCTTTAAGACACAT 600

QY 555 ggaacacacacacggctgagctgacagagccttcacaaaggtctgtgcgataccagcaagt 614
Db 601 GGCACACACACCGCTGAGCTGACAGGCGTTCACCAAGGCGTGTGCGATTCAGCCCAAGGT 660

QY 615 ggtccactgcaagcgtctgtaacctcaagggcggttccacccgagctggaagaaga 674
Db 661 GGTCCACTGTCACACCGCTGAGCTCAAGGGCGCGGCTGCCACCGAGCTGACAAAGCA 720

QY 675 ctctctgctgtgagagcgctgtggagtgtaatgcaagagtgctgcatccccccggagc 734

```

```

Db 721 CTTCTTCTGTGAGAGCGCTGTGGAAAGTGAATGAAGAGATGTGCATCCCCCGGAC 780
QY 735 gtgacctctctgctgagctgacacagagagtgctgtgctgcagcccaagctctgttcaa 794
Db 781 GTTGCTCTCTGCTGGGTGTGCAACGAGAGTGCCTGTGTGAGCCCAAGACTCTGTGTCAA 840
QY 795 ctatggagctgcatgtgtgtgtgtgcaagggcatctctacacatgcagaaatgaagagca 854
Db 841 CTATGAGCAGCTGATGATGTTGTGTGCAAGGGCATCTTACACTGACGATATGAGACGA 900
QY 855 tgaaggtctctgctgacacacccctgctctctcccgctcaactgctgagcgccgctg 914
Db 901 TGAAGGCTCTCTGCGCTACACACCCCTGCTCTCTCCGCTCCAACTGCTGCGCCCGCTG 960
QY 915 gtccctcaatggtgtctctcagtgagtgatgctgctgctgctgtaacctgtgccaac 974
Db 961 GTCTTCATGAGGTGCTCTCTCCGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 975 cggctgtgtgaagctgagccagcgtgtgctagacgcttctgcgcgcgcctgtgtgcgctg 1034
Db 1021 GGGCTGCTGAACCTGTGCGCGAGCGTGGCTACGACCTGTGCGCGCTGTGCTGCGCTG 1080
QY 1035 caagcacacgaagaacgctcactgcaaaagcagcagcgagatgccaagacacagcagcgcc 1094
Db 1081 CAAGCACACGACAGCGTATCTGCAAGAGCAGCGAGCGGAGATGCCAAGCAGCAGCGCC 1140
QY 1095 cgaagagcctcttgcagagttgtgtcgaaagcccgatgctgctgctgaaacctgttc 1154
Db 1141 CGACAAAGCTTTTGTGACAGTTTGTGTGAAAGCCCGAGTCTGCTGCTGCAAACTGTGTT 1200
QY 1155 tctctgacatcaagaagaacgctcagcaaggtcagaggtcttagcctctgagagctgacc 1214
Db 1201 TCTTGTGACATCTAAGAAGACTGACAGAGTCAAGAGTTTAAAGCTCTGAGAGCTGAGC 1260
QY 1215 tgcctagctgtgcccactccctaccccgagcttcggaataacagagacacacacagctga 1274
Db 1261 TTGCTAGTGTGCCACCTCTTACCCCTGCAAAATACAGAGACACACACACAGTA 1320
QY 1275 cccctgattccccaagaatgataagagaagacatttgagcttcttcaagggctcgaagaac 1334
Db 1321 CCTGTATTTCCCAAGGTGATGAAGAGCACTTGGGCTTTTTCAGAGGTCTGTAAC 1380
QY 1335 ttgtgtcaaaacagaacaaagcagagggatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1394
Db 1381 TTGTGTCAAAACAGACAAATGAGAGGCGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
QY 1395 agaaagacagaacacagatgtgagacacataccggaacatgcagctgagctgaatgccttc 1454
Db 1441 AGAAGACAGAACACAGATGTGACACATATCCGGAACATGACAGCTTGAAATGCCCTTC 1500
QY 1455 caagccctctctctccctccctccctccctccctccctccctctctctctctctctctctct 1514
Db 1501 CAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1560
QY 1515 cactcaagagagctagctgctgagaggaattgtctaactgaatccaaaggttaac 1574
Db 1561 CTCTCACAGAGAGCTAGCTGCTGGAGGAATGTTAACGTATACAGAGGTACCTTTAA 1620
QY 1575 gaagaccttgagagctctctataccttctctctctctctctctctctctctctctctctct 1634
Db 1621 GAAGACCTTGGAGTCTTATATACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680
QY 1635 gtccctgagctctgtggaagtgtaga 1662
Db 1681 GTCCCTGATGCTTGTGGGAAGGTGAGA 1708

RESULT 2
AF176906 1405 bp mRNA linear ROD 23-DEC-1999
LOCUS AF176906
DEFINITION Mus musculus sprouty 4 (Spry4) mRNA, complete cds.
ACCESSION AF176906

```


*	569	1350: contig of 782 bp in length	*			*			
*		gap of unknown length	*			*			
*	1351	1578: contig of 228 bp in length	*			*	30200	31063: contig of 864 bp in length	
*		gap of unknown length	*			*		gap of unknown length	
*	1579	2731: contig of 1153 bp in length	*			*	31064	32094: contig of 1031 bp in length	
*		gap of unknown length	*			*		gap of unknown length	
*	2732	3366: contig of 635 bp in length	*			*	32095	32924: contig of 830 bp in length	
*		gap of unknown length	*			*		gap of unknown length	
*	3367	4347: contig of 981 bp in length	*			*	32925	33506: contig of 582 bp in length	
*		gap of unknown length	*			*		gap of unknown length	
*	4348	5253: contig of 906 bp in length	*			*	33507	34428: contig of 922 bp in length	
*		gap of unknown length	*			*		gap of unknown length	
*			*			*	34429	35484: contig of 1055 bp in length	

```

*      61365      62047: contig of 683 bp in length
*      *          gap of unknown length
*      62048      62768: contig of 721 bp in length
*      *          gap of unknown length
*      62769      63630: contig of 862 bp in length
*      *          gap of unknown length
*      63631      64565: contig of 935 bp in length
*      *          gap of unknown length
*      64566      65334: contig of 769 bp in length
*      *          gap of unknown length
*      65335      66251: contig of 917 bp in length
*      *          gap of unknown length
*      66252      66976: contig of 725 bp in length
*      *          gap of unknown length
*      66977      67842: contig of 866 bp in length
*      *          gap of unknown length

```

```

Query Match      49.8%; Score 827.8; DB 2; Length 155614;
Best Local Similarity 99.6%; Pred. No. 3.2e-168;
Matches 840; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

```

QY 604  cagcccaagatggtcctcctgacgacgctgacactcaaggcgaggggtccacccgag 663
      |||||||
Db 39394  CAGCCCAAGATGCTCCTCACTCCAGCCGCTGACCTCAAGGCGCGGCTCCACCCGAG 39395
QY 664  ctgagcaagcaactctctctgctgagagcctgtagaagttaatgcaagagtgagca 723
      |||||||
Db 39334  CTGAGCAAGCACTCTCTCTCTGCGAGGCGCTGAGAACTGAATGCAAGAGCTGTGCA 39275
QY 724  tccccccgagaggtgctcctcctctgctgagctgcaacgaagagtgctgctcagccag 783
      |||||||
Db 39274  TCCCCCGGAGGTTGCTCTCTGCTGGCTGCAACGAGAGTCTGCTCAGCCAG 39215
QY 784  actctgcaactatgagcagtgatgattgtgtgtagggcagctctctacactgagc 843
      |||||||
Db 39214  ACTCTGCACTATGAGACACTGATGTTGTGTGTCAGGCGCATCTCTACCTGACG 39155
QY 844  aatgagagcagtgaggggtcctgctgctgagcaccctgctcctcctccgctcccaactgc 903
      |||||||
Db 39154  AATGAGAGCATGAGGGCTCTCTGCGCTGACCA -CCCTGCTCTCTCCCGCTCCACTGCG 39096
QY 904  tggcgccgctgctgctcctcctcctgctgctcctcctgctgctgctgctgctgctgctgct 963
      |||||||
Db 39096  TGGCGCCGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 39036
QY 964  ctgctcctgcaacgagctgctgctgagcctgagcctgctgctgctgctgctgctgctgctgct 1023
      |||||||
Db 39036  CTGCTCTGCAACGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 38976
QY 1024  ggttgccgctgctgcaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1083
      |||||||
Db 38976  GGTGCGCCTGCAACGACACGACAGCGCTCATCTGCAAAAGCAGCGCGGAGATGCGCAAG 38916
QY 1084  accagcagggccgagcaagcctctctgagcagctgctgctgagcagcagcagcagcagcagcag 1143
      |||||||
Db 38916  ACCAGCAGGCGCGCAAGCCTCTCTGAGAGTTGTGTCGAAGCCCGAGTCTGCTGCTGCG 38856
QY 1144  aaactctgctcctctctgacacacaaagaaagactgacagagtgctgagctgctgctgctgct 1203
      |||||||
Db 38856  AAACCTGGTTCTCTCTGACATCTAAGAAGACTGCACAGGTGAGAGGTTTAAAGCTTCC 38796
QY 1204  tgaagctgacctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1263
      |||||||
Db 38796  TGAGGCTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 38736
QY 1264  acccagcagtaacctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1323
      |||||||
Db 38736  ACCCAGCAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 38676
QY 1324  ggtctcgaactctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1383
      |||||||
Db 38676  GGTCTGAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 38616

```

```

QY 1384  ttctcttcagaagacagacagatgtgacacatctccggaaactgctgctgct 1443
      |||||||
Db 38615  TTTCTTTTTCAGAGACAGACAGATGTGACACATATCCGGAACATGCACTGCTTT 38556
QY 1444  gaa 1446
      ||
Db 38555  AAA 38553

```

```

RESULT 5
AC099211 183042 bp DNA linear HTG 20-DEC-2001
LOCUS Rattus norvegicus chromosome Ch18 clone CH230-34E17, ***
DEFINITION SEQUENCING IN PROGRESS ***, 79 unordered pieces.
ACCESSION AC099211
VERSION AC099211.2 GI:11973092
KEYWORDS HTG; HTGS; PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 183042)

```

REFERENCE

AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., All-ouman,F.R., Allen,C., Alsbrooks,S.L., Amaralungu,H.C., Are,J.R., Banks,T., Barbarta,J., Benton,J., Blmage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Dwyer-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Dera,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K., Harris,K., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Hollway,C., Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,D., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louissege,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mathoney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokemko,S., Ogih,M., Okunou,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Prins,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojudoan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Slisnon,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stinson,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,C., Telford,B., Thomas,R., Thomas,S., Usmani,K., Vasquez,L., Vera,Y., Villalón,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S., Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G., and Gibbs,R.

TITLE

JOURNAL

2 (bases 1 to 183042)

AUTHORS

JOURNAL

Submitted (09-NOV-2001)

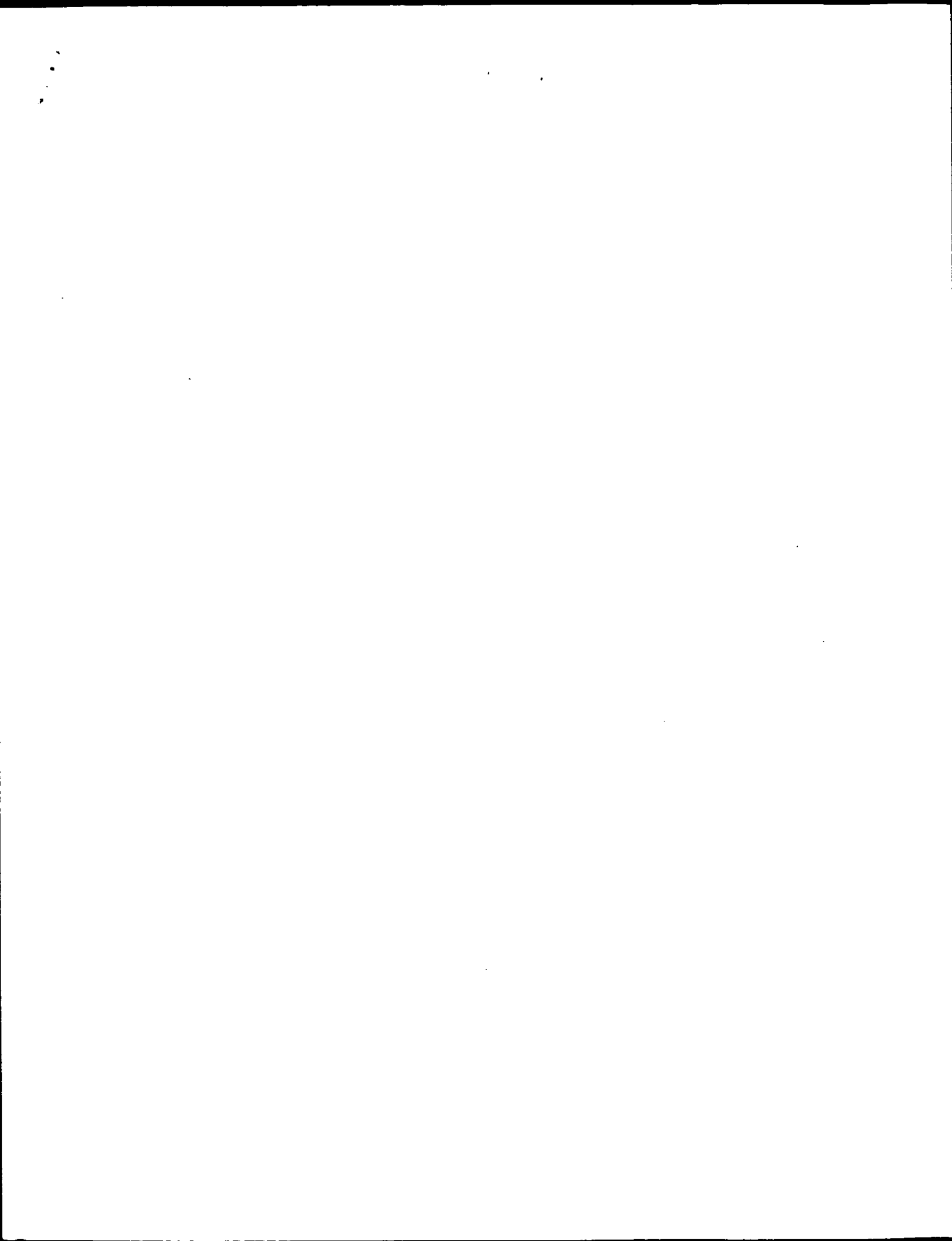
COMMENT

On Dec 20, 2001 this sequence version replaced gi:16874705.

Center: Baylor College of Medicine

QY 138 gaccccaagcccaatgagagcccccagatcccaaga---gcccgcctttagctcccaactc 254
 DB 15953 GACCTCCAGCCATGAGAGCCGCCCATTCACAGAGCGCGTCCCTTGAATCCAGCTC 16012
 QY 255 agctatgctcaagccctctctttagacagcgatgtcccaagcgctctcaagcccaact 314
 DB 16013 AGTCATGGTGAGCCCTCTCTTGAACAGAGAGTCCCAAGCGCGCTCCAGACCCACT 16072
 QY 315 caccatctcaagccatctgagcagcgatgaagaccagcaatgttgagaaatgtactatagaaca 374
 DB 16073 CACCATCTGGCCATTCAGATGACAGATGAAGACAGCCAGCTGAGAAATGACTACATAGACAA 16132
 QY 375 cccatagcttgagcccttgaacagcgagcccaagcgagccggcgagccagagagctggc 434
 DB 16133 CCCTAGCCCTGGGCCCTGTGCACAGGGTCCCAAGCGCGGTGGGGGCCCCAGAGCTGGC 16192
 QY 435 cccgagcccgccgcgctgtgacagagatgtcaaccacatctgacatctctcaagcgagcg 494
 DB 16193 TCCTACACCTGCGCCCTGTGACAGAGATGTACTACACAGATCTCTTCCAGCGTGC 16252
 QY 495 cccagcctctgtagcagcagcagagacacatctctgacaaagcgctcttagaacacat 554
 DB 16253 GCCAGCTCTGTGATAGCAGAGAGAGAGACACTTCTGTGACAGCGGCTCTGTGATACAT 16312
 QY 555 ggcacacacacccgctgtgacagcagcagcagcagcagcagcagcagcagcagcagcag 614
 DB 16313 GGCCTCCACACCGGTGACTGAGAGAGCGCTCCCGAGGGCTGTGCGCTCCAGCCCAAGT 16372
 QY 615 ggtcacctgacagcgctgtgacacacagcgagcgagcgagctccacacagcagcagcagc 674
 DB 16373 GGTCCACATGTAAAGCCACAGGAGCTCAAGGCGCCAGACCTCCACAGAGCTAGCAAGCA 16432
 QY 675 ctctctgtgtgagagcctgtgagaaatgttaaatgcaagagatgtgatacccccgagc 734
 DB 16433 CTCTCTGCTGTGCGAGCGCTGTGGAAATGTAAATGCAAGAGAGTGTGCTCCCTCCGAC 16492
 QY 735 gtcgctctctgt 794
 DB 16493 GTTACCCCTGCTGT 16552
 QY 795 ctatgagcagctgt 854
 DB 16553 CTACAGCAGTGTGATGT 16612
 QY 855 tgaaggtctctgt 914
 DB 16613 TGAGGGCTCTGT 16672
 QY 915 gtctctctgt 974
 DB 16673 GTCTCTTGT 16732
 QY 975 cggctgt 1034
 DB 16733 GGGCTGCTCAAGCTGT 16792
 QY 1035 caagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1094
 DB 16793 CAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 16852
 QY 1095 cgaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1154
 DB 16853 CGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 16888
 QY 1155 tctctgacatctaaagagactgtgagaaagtttagcctctcagagctgac 1214
 DB 16889 -----GTCCCTCTGCGACATCTCTGAGGCTGAC 16918
 QY 1215 ttgctatctgacacactccctacccagcagcttgcgaataa---cagagacccaccca 1270
 DB 16919 TTGCTCATCTGTCT 16978

QY 1271 cgtaccctgtatctcccaagatgaaagacacttgggctttttcagggctctg 1330
 DB 16979 CTTGACCTCGAGTCCCAAAAGATGAGATACCTTTGGGG---GTTTCATGTGCTTG 17035
 QY 1331 aaacttgtgtcaaacagacaaatgcaagcgaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1390
 DB 17036 AAACCTGTGTAACAGCAGTGTGACAGTGTGCGGAA-----GTGTAAGGGGCGAGTTTCTTA 17090
 QY 1391 ttcaagaagacagaaacagatgtgacacatctcggaactcgagcgtctgaaatgac 1450
 DB 17091 TTTCACAAAGATGTAAAGACAGTGTGACACAAATCTGCAAGTTTCACACCTCAATGTGCC 17150
 QY 1451 ttccagccct 1498
 DB 17151 TTCCAGCCCT 17197
 RESULT 6
 AF227517 7053 bp mRNA linear PRI 02-FBB-2001
 LOCUS AF227517
 DEFINITION Homo sapiens sproutly-4C mRNA, complete cds.
 ACCESSION AF227517
 VERSION AF227517.1 GI:12655914
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 7053)
 Leeksa, O.C., van Achterberg, T.A.E., Spaargaren, M., Kr. von dem
 Borne, A.E.G., Pannekoek, H. and de Vries, C.J.M.
 Identification of a novel human sproutly homolog
 unpublished
 TITLE JOURNAL
 REFERENCE 2 (bases 1 to 7053)
 de Vries, C.J.M., Leeksa, O.C., van Achterberg, T.A.E. and
 Pannekoek, H.
 Direct Submision
 Submitted (21-JAN-2000) Biochemistry, Academic Medical Center,
 Meibergdreef 15, Amsterdam 1105 AZ, The Netherlands
 FEATURES
 source
 1..7053
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_type="smooth muscle"
 /tissue_type="umbilical artery"
 206..526
 /codon_start=1
 /product="sproutly-4C"
 /protein_id="AAK00653.1"
 /db_xref="GI:12655915"
 /translation="MEPPIPOSAPIPLPNSVMYOPILDSMSHSRLQHPILPLIDOVK
 TSHVENDYIDNLSLALTGPKRTRGAPLAPTRRCQDVTYHHMISGRCSATCL
 PPA"
 CDS
 BASE COUNT 1696 a 1740 c 1746 g 1871 t
 ORIGIN
 Query Match 44.2% Score 734.2 DB 9: Length 7053;
 Best local Similarity 72.4% Pred. No. 5.2e-148;
 Matches 1200; Conservative 0; Mismatches 3; Indels 454; Gaps 1;
 DB 1 GCGAGCTGAGACTACAGCGCGAGCTGTGGCTGTGAGGAGGAGCGCTTGCGGGTTC 60
 QY 6 gcgagctgagctgacagcgagcgt 65
 DB 1 GCGAGCTGAGACTACAGCGCGAGCTGTGGCTGTGAGGAGGAGCGCTTGCGGGTTC 60
 QY 66 tccgacggcgctgtgagtagacagcgagcgttaacctgcccggctctcagatattacaca 125
 DB 61 TTCGACCGGGGCTGTGAGTACAGCGCGGCTTAACCTGCGCGGCTTCAGGATTTACACA 120
 QY 126 gacgtggcgagatgt 185
 DB 121 GACGTGGGCGATGCTTGTGACCTCTCAAGTCTCTCAAGGCCCTTGAAGGCTGTTCTC 180



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:46:18 ; Search time 24.31 Seconds
(without alignments)
1799.071 Million cell updates/sec

Title: us-10-082-902-2
Perfect score: 1657
Sequence: 1 MEPPITQSAPLTPNSVMWQP.....SVICKAASGAKTSRDRPF 299

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1657	100.0	322	4 Q9C004	Q9C004 homo sapien
2	520	31.4	106	4 Q9C003	Q9C003 homo sapien
3	154.5	9.3	409	5 Q9V756	Q9V756 drosophila
4	126	7.38	11	Q88424	Q88424 mus musculus
5	125.5	7.6	984	5 Q9Y1P7	Q9Y1P7 cryptospori
6	121.5	7.3	757	5 Q9V2F2	Q9V2F2 drosophila
7	119.5	7.2	1365	4 Q96028	Q96028 homo sapien
8	116	7.0	1187	2 Q49549	Q49549 mycoplasma
9	112.5	6.8	470	4 Q9Y4Z3	Q9Y4Z3 homo sapien
10	112.5	6.8	617	4 Q9Y4Z4	Q9Y4Z4 homo sapien
11	111.5	6.7	723	11 Q9JWS4	Q9JWS4 mus musculus
12	111.5	6.7	723	11 Q9VWH2	Q9VWH2 mus musculus
13	111.5	6.7	734	4 Q9Y112	Q9Y112 homo sapien
14	111.5	6.7	832	4 Q9C010	Q9C010 homo sapien
15	111	6.7	667	5 Q9XTK3	Q9XTK3 giardia lam
16	109	6.6	432	4 Q9NPM2	Q9NPM2 homo sapien
17	109	6.6	494	4 Q9S965	Q9S965 homo sapien
18	109	6.6	898	4 Q9UF24	Q9UF24 homo sapien
19	109	6.6	4599	11 Q9J118	Q9J118 mus musculus

20	108.5	6.5	661	11 Q9JKS3	Q9Jks3 mus musculus
21	108	6.5	495	5 Q9GQ43	Q9Gq43 giardia lam
22	108	6.5	4599	4 Q9NZR2	Q9nZR2 homo sapien
23	107.5	6.5	4123	4 Q75851	Q75851 homo sapien
24	107	6.5	432	5 Q22884	Q22884 caenorhabdi
25	107	6.5	496	5 Q9VMD4	Q9vmd4 drosophila
26	107	6.5	4823	13 Q93321	Q93321 fugu rubrip
27	106.5	6.4	1119	5 Q18034	Q18034 caenorhabdi
28	106.5	6.4	1513	5 Q17970	Q17970 caenorhabdi
29	106.5	6.4	1616	5 Q9V5J0	Q9vsj0 drosophila
30	106	6.4	721	13 Q91902	Q91902 xenopus lae
31	106	6.4	5374	11 Q99ND0	Q99nd0 mus musculus
32	105.5	6.4	1704	5 Q94446	Q94446 chironomus
33	105.5	6.4	1713	11 Q88349	Q88349 mus musculus
34	105.5	6.4	3857	11 Q88840	Q88840 mus musculus
35	105	6.3	1534	4 Q75093	Q75093 homo sapien
36	105	6.3	2146	5 Q9VC97	Q9vc97 drosophila
37	104.5	6.3	1587	4 Q9Y5N6	Q9y5n6 homo sapien
38	104	6.3	403	5 Q18375	Q18375 drosophila
39	104	6.3	664	4 Q9U1L7	Q9u1l7 homo sapien
40	104	6.3	1664	5 Q9TVQ2	Q9tvq2 caenorhabdi
41	103.5	6.2	806	5 P91808	P91808 strongyloce
42	103	6.2	734	10 Q9LNZ0	Q9lnz0 arabidopsis
43	103	6.2	1420	5 Q9GZJ3	Q9gzj3 leishmania
44	103	6.2	1511	4 Q75412	Q75412 homo sapien
45	103	6.2	2872	11 Q9WUH8	Q9wuh8 rattus norv

ALIGNMENTS

RESULT 1

Q9C004 PRELIMINARY; PRT; 322 AA.
AC Q9C004;
DF 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE SPROUTY-4A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-UMBILICAL ARTERY;
RA Leeksa O.C., van Achterberg T.A.E., Spaargaren M.,
Kr von dem Borne A.E.G., Pannekoek H., de Vries C.J.M.;
RT "Identification of a novel human Sprouty homolog.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF227516; AAK00652.1;
SQ SEQUENCE 322 AA; 34929 MW; 5BF5B67854CDD6DD CRC64;

Query Match	100.0%;	Score	1657;	DB	4;	Length	322;
Best Local Similarity	100.0%;	Pred. NO.	1.2e-142;				
Matches	299;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MEPPITQSAPLTPNSVMWQP	LDLSRMSHSLQHP	LTILP	LDQVKTSHVENDY	IDNP	LSAL 60
Db	24	MEPPITQSAPLTPNSVMWQP	LDLSRMSHSLQHP	LTILP	LDQVKTSHVENDY	IDNP	LSAL 83
Qy	61	TTGPKRRGGAP	ELAPAPARCDQVTHWISFSGRPSVSSSSSTSSDQRLLDHMAPPV				120
Db	84	TTGPKRRGGAP	ELAPAPARCDQVTHWISFSGRPSVSSSSSTSSDQRLLDHMAPPV				143
Qy	121	ADQASPRAVRIQPVVH	COPLDLKGPVPP	ELDKHFL	LLCEACGKCKE	CAKES	PTLPSCW 180
Db	144	ADQASPRAVRIQPVVH	COPLDLKGPVPP	ELDKHFL	LLCEACGKCKE	CAKES	PTLPSCW 203
Qy	181	VCNOECLCSAOTLVNYGTCM	CLVOGIFYHCTN	DEDEGSCADHPC	SCSRNCCAR	WSPMA	240
Db	204	VCNOECLCSAOTLVNYGTCM	CLVOGIFYHCTN	DEDEGSCADHPC	SCSRNCCAR	WSPMA	263

```

QY 241 LSVWLPCLCYLPATGCVKLAQGYDRLRRPCCRCKHTNSVICAAAGDARTSRPKPF 299
DB 264 LSVWLPCLCYLPATGCVKLAQGYDRLRRPCCRCKHTNSVICAAAGDARTSRPKPF 322

RESULT 2
Q9C003 PRELIMINARY: PRT: 106 AA.
AC Q9C003;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE SPROUTY-4C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=UMBILICAL ARTERY;
RA Leeksa O.C., van Achterberg T.A.E., Spaargaren M.,
RA Kr von dem Borne A.E.G., Pennakoeck H., de Vries C.J.M.;
RT Identification of a novel human Sprouty homolog.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF227517; AAK00653.1; -.
SQ SEQUENCE 106 AA; 11482 MW; 6158571E3477A598 CRC64;

Query Match 31.4%; Score 520; DB 4; Length 106;
Best Local Similarity 97.0%; Pred. No. 6.le-40;
Matches 97; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MEPPIPQAPLTPNSVMVQPLDLSRMSHSRLQHPILTPIDQVKTSHVENDYIDNPSIAL 60
DB 1 MEPPIPQAPLTPNSVMVQPLDLSRMSHSRLQHPILTPIDQVKTSHVENDYIDNPSIAL 60

QY 61 TTGPKRTGGAPLAPTPARCDQDQVTHHWISFGRPSVS 100
DB 61 TTGPKRTGGAPLAPTPARCDQDQVTHHWISFGRPCSAT 100

RESULT 3
Q9V756 PRELIMINARY: PRT: 409 AA.
AC Q9V756;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CG10155 PROTEIN.
GN CG10155.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Branton R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

```

```

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler K., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan W., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003813; AAF58207.1; -.
DR FlyBase; FBgn0033964; CG10155.
DR InterPro; IPR000697; RnBp1_WASP.
DR SMART; SM00461; WH1; 1.
SQ SEQUENCE 409 AA; 44311 MW; 6FC0BD89FF3F9C96 CRC64;

Query Match 9.3%; Score 154.5; DB 5; Length 409;
Best Local Similarity 23.2%; Pred. No. 3.2e-06;
Matches 66; Conservative 22; Mismatches 99; Indels 97; Gaps 14;

QY 52 YIDNPSIALTTGPKRTGGAP--ELAPTPARCDQ-----DVTHHWISF----- 92
DB 176 YISTDKTSATSTPP-----DAPPASAAAPSPATAGIAAENSYVTLTAVHHYNYPPVDQ 231

QY 93 -----SGRPSSVSSSSSTSSDQRLDHPAPPVADQASPRAVRQPKVHCQP----L 141
DB 232 PVGAQVNLNARRESISALKRNALE-----AAQMAAAQTAAAGGLACRDGSGK 280

QY 142 DLKGPVPPPELDKHF-LLCEAC-----GKCKCEKASPRTPIPSCWVCNQECICS 189
DB 281 PLHKPNVSDILKKETRLRCRYCHELYSEDFNRRGACE-----YAPDAFRSGYECIS- 331

QY 190 AQTLYVNGTCMLVQGIIFYHCTHEDDEGSCADHPGCSRS--NCCARWSPMGALSVP 247
DB 332 -----GNGCARCMI-----YHCMS--DAEGETAQHPDCSASEAGCSRWLGATLSLFVPC 381

QY 248 LLCLYPATGCVKLAQGYDRLRRPCCRCKHTNSVICAAAGDARTSRPKPF 291
DB 382 LWCPPLPRLAC-----HLAGHGLCGGQHK 406

RESULT 4
O88424 PRELIMINARY: PRT: 738 AA.
ID O88424;
AC O88424;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PACTOLUS.
GN ITGB2L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-NIH(S); TISSUE-BONE MARROW;
RX MEDLINE-98204856; PubMed-9535848;
RA Chen Y., Garrison S., Weis J.J., Weis J.H.;
RT *Identification of pectolus, an integrin beta subunit-like cell-
RT surface protein preferentially expressed by cells of the bone
RT marrow.*;
RL J. Biol. Chem. 273:8711-8718(1998).
CC -!- FUNCTION: INTEGRINS ARE A LARGE FAMILY OF CELL SURFACE
CC GLYCOPROTEINS THAT MEDIATE CELL TO CELL & CELL TO MATRIX ADHESION.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
CC BONDS (BY SIMILARITY).
CC -!- SIMILARITY: WITH OTHER BETA CHAINS FROM THE INTEGRIN FAMILY OF
CC CELL-SURFACE RECEPTOR.
DR EMBL; AF051367; AAC25502.1; -.
DR MGI; 1277979; Itgb2l.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002369; Integrin_B.
DR InterPro; IPR001169; Integrin_beta_C.
DR InterPro; IPR003659; PSI.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF00362; Integrin_B; 2.
DR ProDom; PD001811; Integrin_B; 1.
DR SMART; SM00187; INB; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00243; INTEGRIN_BETA; 3.
KW Cell adhesion; Cytoskeleton; EGF-like domain; Extracellular matrix;
KW Glycoprotein; Integrin; Repeat; Transmembrane.
SQ SEQUENCE 738 AA; 81605 MW; E6B67A1554C6FIDA CRC64;

Query Match 7.6%; Score 126; DB 11; Length 738;
Best Local Similarity 21.0%; Pred. No. 0.0021;
Matches 78; Conservative 40; Mismatches 108; Indels 146; Gaps 21;

QY 22 LDRMSHRLQHPILPILPDQVKTSHVENDYIDNPSLALTTGPKRTGGAPELAPTPARC 81
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
279 LASKLAENNIQ-PIEVPSPRMVKTVEKLTTFI-----PKLTIG---ELSDSSNV 324

QY 82 DQDVTHHVISFG-----RPSSVSS-----SSSTSSDQRLLDHMAPPPVADQASPR 127
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
325 AOLIRNAYSKLSSIVLHNSHTIPSLKYTDYSYCSNGTSNPKPSGDCSGVINDQVTFQ 384

QY 128 -----AVRIQPKVYHCOPIDLKGPVAPPPELKHFLICEA- 161
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
385 VNITASECFREQFFIQLGFMDSVTVRLP-LCECQ-----COEQSOHHSLCGGK 434

QY 162 -----CGKCKC-----KCAPRTLP-----SCWVCNQCCLSAQTLVNYGTCM- 201
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
435 GAMEGICRCNSGYAGKNCCEQQTQGPSSQDLGGCKDNSSIMCS-----GLGDCICQC 489

QY 202 -----LVQGYFYHCTN---EDDEGSCADHP-----CSCSRNCCARWFMGALSV 244
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
490 ECHTSDIPNKEIYGYCECDNVNCRYDQVCGGPERGHCSCGR--CFCRYSFVGS----- 543

QY 245 LPCLLCYLPATGCVK-----LAQRGYDR---LRRPG-----C 273
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
544 --ACQCRSTSGCLNRMVVEGSGHGRVCNRCCLDPGYQPLCEKRPGRFYHRCSEYYS 601

QY 274 RCKHTNSVI-CK 284
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
602 RCLKDNSAIKCR 613

RESULT 5
ID Q9Y1P7 PRELIMINARY; PRT; 984 AA.
AC Q9Y1P7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

```

```

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE HYPOTHETICAL 106.3 KDA PROTEIN.
GN HC-10.
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=3807;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99236038; PubMed-10219298;
RA Schroeder A.A.; Lawrence C.E.; Abrahamson M.S.;
RT "Differential mRNA display cloning and characterization of a
RT Cryptosporidium parvum gene expressed during intracellular
RT development.*;
RL J. Parasitol. 85:2213-2220(1999).
DR EMBL; AF097183; AAD42044.1; -.
DR HSP; P04355; 2MRT.
KW Hypothetical protein.
SQ SEQUENCE 984 AA; 106294 MW; 5860CC129BB26120 CRC64;

Query Match 7.6%; Score 125.5; DB 5; Length 984;
Best Local Similarity 21.9%; Pred. No. 0.0031;
Matches 84; Conservative 39; Mismatches 155; Indels 105; Gaps 21;

QY 8 SAPLTPNSVMVQPLDLSRMSHSRLQHPILPILPDQVKT-----SHVENDYIDNP-----SL 58
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
379 SPVVTNPKRLAEATSDIEEDDEMPNVOETILD-PQVSTPEILSKVATGKAYTTPAQACASI 437

QY 59 ALTTGPK--RTRGGAPELAPTPARCDDQVH---HWTS----- 91
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
438 SCLSGPSCISNKGSPKCYKATCVKKNHIAIHVDCSDAKCEKSCGKNCVCGECK 497

QY 92 -----FSGRPSSVSSSSSTSDQRLLDHMAPPP--VADQASPRVR-----IOPKVV 136
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
498 PGKCVGSGSHGSSSTKCVGCTTEVPRLGCVARPAHCTGGSCSPARVARGCVKHEIL 557

QY 137 ---HCQPLDL-KGPAVPELKHFLICEAGC-KCKCKEASPRPLP-SC-----WVCN 183
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
558 PNHCAKEGSKGCVKGECKKGVAGDCSHRCYCECVAPRCQPGCTGTCTTGCN 617

QY 184 QECCLCSAQLVNYGTC---MCL---VOGIFY-----HCTNEDDEGSCADHPSCSRS 229
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
618 ---VGESQTGCEKSGKGGYCVGPKCNGNEYDGFCKGSECTPAKCHPSC--NGCSCTTA 672

QY 230 NC-----CARWFMGALSVMPLCLLCYLPATGCVKLAQRYDRLRRPQCRC 275
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
276 KHTNSVICKAASGDAKTSRPDKP 298
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
726 GYDGRAVGSNNYKSHCAGKP 748

RESULT 6
ID Q9VZF2 PRELIMINARY; PRT; 757 AA.
AC Q9VZF2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CG15011 PROTEIN.
GN CG15011.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D.; Celniker S.E.; Holt R.A.; Evans C.A.; Gocayne J.D.,

```

RA Ananaitides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George K.A., Lewis S.E., Richards S., Ashbourne M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Beso P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Feiraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of *Drosophila melanogaster*.";
 RT Science 287:2185-2195(2000).
 RL EMBL: AE003480; AAF4781.1;
 DR FlyBase; FBgn0035518; CG15011.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001368; TNFR.C6.
 DR InterPro; IPR000967; ZnF.NFX1.
 DR InterPro; IPR001841; ZnF.ring.
 DR Pfam; PF01422; zf-NFX-1; 7.
 DR SMART; SM00184; RING; 1.
 DR SMART; SM00438; ZnF.NFX; 8.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
 SQ SEQUENCE 757 AA; 84556 MW; 34A7F3C28BD6428F CRC64;

Query Match 7.3%; Score 121.5; DB 5; Length 757;
 Best Local Similarity 24.6%; Pred. No. 0.0056;
 Matches 70; Conservative 16; Mismatches 103; Indels 95; Gaps 18;

QY 64 PKRTRGGAPLAPPPARCDQDQVTHHWFSGR-----PSSVSS-----SSSTS 106
 Db 148 PKCRDRYQPADKPTQYNC-----FCGKEVNPENQPLVPHSGEHCGLLPKCG 197
 QY 107 SDQRLDHRMAP-PPVAQQA-----SPRAVRIQKVHVHCOPLDLKGPVAPPDLKHF 156
 Db 198 HDCKLLCHPGPCPCAQAQVSCLCGKSSPSRVCIDKQWTCQ-----QTCKEL 246
 QY 157 LLCBACGKCKKE-CASPTRLPSGCWNCN-QECLCSAQT-LVNYGTCMLVQGI----- 206
 Db 247 L---ACGKHKCNQVCHQPGKPCPTSKLQPCQECQRESKMNVCNSDRKWKCONVCGAPFAC 303
 QY 207 -FYHCTNDDGGCAD-----HPCSCSRNSCCARSPMGALSIVLPC-----LLCY 251
 Db 304 GLHICEKVCVHAGPGDGECPQLQVNSCPGK-NCQG-----DQCPPECIKCGKQLSC- 353
 QY 252 LPATGCVKLAQRGYDRLRRRGCRCKHFNVSICRAASGDAKTSRP 295

Db 354 -NKHKQSVCHNG-----PCYPCKLESQINCRG--GKTKKSV 388
 RESULT 7
 Q96028 PRELIMINARY; PRT; 1365 AA.
 ID O96028
 AC O96028;
 DT 01-MAY-1999 (TEMBLrel. 10, Created)
 DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
 DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
 DE WHSCL PROTEIN;
 GN WHSCL OR MMSET OR TRX5.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98482232; PubMed=9618163;
 RA Stec I., Wright T.J., Van Ommen G.J.B., de Boer P.A.J.,
 RA van Haeringen A., Moorman A.F.M., Altherr M.R., Den Dunnen J.T.;
 RT "WHSCL, a 90 kb SET domain-containing gene, expressed in early
 development and homologous to a *Drosophila* dysmorph gene maps in the
 RT Wolf-Hirschhorn syndrome critical region and is fused to IgH in
 RT t(4;14) multiple myeloma.";
 RL Hum. Mol. Genet. 7:1071-1082(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX TISSUE=MYELOMA;
 RX MEDLINE=99005326; PubMed=9787135;
 RA Chesni M., Nardini E., Lim R.S., Smith K.D., Kuehl W.M.,
 RA Bergsagel P.L.;
 RT "The t(4;14) translocation in myeloma dysregulates both FGFR3 and a
 RT novel gene, MMSET, resulting in IgH/MMSET hybrid transcripts.";
 RL Blood 92:3025-3034(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Angrand P.O., Valvatne H., Jeannoulin F., Adamson A.,
 RA van der Hoeven F., Olsen L., Tekotte H., Huang N., Poch O.,
 RA Lamerding J., Chambon P., Lossen R., Stewart A., Aasland R.;
 RT "Mammalian Trithorax- and ASH1-like proteins: Putative chromatin
 RT regulators which contain PHD fingers and SET domains.";
 RL Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Chesni M., Bergsagel P.L.;
 RA Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF083386; AAD19343.1;
 DR EMBL: AF071593; AAC24150.1;
 DR EMBL: AJ007042; CAB45386.1;
 DR EMBL: AF083387; AAD21770.1;
 DR EMBL: AF083388; AAD21771.1;
 DR EMBL: AF178219; AAF23370.1;
 DR EMBL: AF178198; AAF23370.1; JOINED.
 DR EMBL: AF178199; AAF23370.1; JOINED.
 DR EMBL: AF178200; AAF23370.1; JOINED.
 DR EMBL: AF178201; AAF23370.1; JOINED.
 DR EMBL: AF178202; AAF23370.1; JOINED.
 DR EMBL: AF178203; AAF23370.1; JOINED.
 DR EMBL: AF178204; AAF23370.1; JOINED.
 DR EMBL: AF178205; AAF23370.1; JOINED.
 DR EMBL: AF178207; AAF23370.1; JOINED.
 DR EMBL: AF178208; AAF23370.1; JOINED.
 DR EMBL: AF178209; AAF23370.1; JOINED.
 DR EMBL: AF178210; AAF23370.1; JOINED.
 DR EMBL: AF178211; AAF23370.1; JOINED.
 DR EMBL: AF178212; AAF23370.1; JOINED.
 DR EMBL: AF178213; AAF23370.1; JOINED.
 DR EMBL: AF178214; AAF23370.1; JOINED.
 DR EMBL: AF178215; AAF23370.1; JOINED.
 DR EMBL: AF178216; AAF23370.1; JOINED.
 DR EMBL: AF178217; AAF23370.1; JOINED.
 DR EMBL: AF178218; AAF23370.1; JOINED.

Query Match	7.0%	Score	116	DB	2	Length	1187
Best Local Similarity	21.9%	Pred	No. 0.027				
Matches	59	Conservative	25	Mismatches	113	Indels	72
						Gaps	14
QY	35	LITLPIDQKTSVENDYIDNPSLALTTGPKRTGAPELAPPARD--QDYTHMISF	92				
Db	266	LTDLEDEHETVHLEEV---CLA-----CHVACDICKNKNEHELLY	306				

Query Match	Similarity	6.8%	Score 112.5	DB 4	Length 470
Best Local	Similarity	21.3%	Score No. 0.023		
Matches	77	Conservative	42	Mismatches	132
				Indels	111
				Gaps	19
Oy	4	PTPQSAFLPNV---	MWPLDSSRSHRLQ---	PTLPLPIDQKSHENDINDP	56
	
Db	134	PAKPRVYTTASTRPSVQV	PPASTYSPSPGAYSTPTTPXDA	-----TTPSP	184
	
Oy	57	SLATTTGPRRTGGABELA	PTPARCDDQVTHHMTSFGSPSSVS	-----	101
	
Db	185	AAAYTPSPVPTVTPSPAP	TPSPAPNTNPAASVAASGGSPAEPAESPPWTDTSFQKA		244
	

QY 102 --SSSTSSDQRLDLMHAP-----PPVADQASPAVR-----OPRVHCQPLDLKG 145
 Db 245 PGKSTSTISKQTLPRGPRATPAAGPVPPLARCTVQRAEREPASSRPLCGHCNNV-IRG 303
 QY 146 PAV-----PELDKHFLLCEACGCK-----CCECASPRLLPSCVW 181
 Db 304 PFLVAMGRSMHPEEP-----TCAYCKTSLADVCXEVEONNVYCEKCYEQFPAPLCAK 355
 QY 182 CNOECILCSAQTLYN---YGTC-MCLV-----OGIFYHCTNEDDEGSCADHPSCSRSNC 231
 Db 356 CNTKINGEVMHALLRQTMHTFCFVCAACKRPFNSLFH---MEDGEPCYCKDYINLFSTKC 412
 QY 232 --C-----ARNSFMGAL-----SVLPCLLCYLPATGCVKLAQRGYDLRRPGCRCKHTN 279
 Db 413 HGCDPVEAGDKFTALGHTWHDTCYICAVCHVNLG-----OPFYSKDRPLCK-KHAH 466
 QY 280 SV 281
 Db 467 TI 468

RESULT 10

09Y424

PRELIMINARY; PRT; 617 AA.

AC 09Y424;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE ZASP PROTEIN (FRAGMENT).
 GN ZASP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SKETLEAL MUSCLE AND HEART;
 RA Faulkner G., Pallavicini A., Formentin E., Comelli A., Ievolella C.,
 RA Trevisan S., Bortoletto G., Scannapieco P., Salamon M., Mouly G.,
 RA Valle G., Lanfranchi G.,
 RT "ZASP, a new 2-band alternatively spliced PDZ-motif protein.",
 RL J. Cell Biol. 0:0-0(0).
 CC -1- SIMILARITY: CONTAINS LIM DOMAIN(S). THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.
 DR EMBL: AJ133767; CAB6728.1; -;
 DR Interpro: IPR000345; Cytc_heme_bind.
 DR Interpro: IPR001781; LIM.
 DR Interpro: IPR001478; PDZ.
 DR Interpro: IPR002965; P-rich_extensn.
 DR Pfam: PF00412; LIM; 3.
 DR Pfam: PF00595; PDZ; 1.
 DR PRINTS: PRO1217; PRICHEXTENS.
 DR PRODOM: PDO00094; LIM; 3.
 DR SMART: SM00132; LIM; 3.
 DR SMART: SM00228; PDZ; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS00478; LIM_DOMAIN_1; UNKNOWN_2.
 DR PROSITE: PS50023; LIM_DOMAIN_2; 3.
 DR PROSITE: PS50106; PDZ; 1.
 DR PROSITE: PS50106; PDZ; 1.
 KM LIM motif: Metal-binding; zinc.
 FT NON_TER 1
 SQ SEQUENCE 617 AA; 66518 MW; 22FC7402352A340 CRC64;

Query Match

Best Local Similarity 21.3%; Score 112.5; DB 4; Length 617;
 Matches 77; Conservative 42; Mismatches 132; Indels 111; Gaps 19;

QY 4 PIPQSAPLTGNSV---MVQPLDLSRMSHSLQH---PLTILPLDQKTSHEVNDIUNP 56
 Db 281 PAKPRPVYVTAIRPSYGVVPASTYSPSGANYSPTPTPSAPA-----YTPSP 331
 QY 57 SLATTTGPKRTGGAPELAPTPARCDDVTHHMSFGSRSSVS----- 101

Db 332 APAYTSPVPTTTPSPAPATPSPAPNYPAPSVAVSGGAEPASRPWTTDSFSQKFA 391
 QY 102 --SSSTSSDQRLDLMHAP-----PPVADQASPAVR-----OPRVHCQPLDLKG 145
 Db 392 PGKSTSTISKQTLPRGPRATPAAGPVPPLARCTVQRAEREPASSRPLCGHCNNV-IRG 450
 QY 146 PAV-----PELDKHFLLCEACGCK-----CCECASPRLLPSCVW 181
 Db 451 PFLVAMGRSMHPEEP-----TCAYCKTSLADVCXEVEONNVYCEKCYEQFPAPLCAK 502
 QY 182 CNOECILCSAQTLYN---YGTC-MCLV-----OGIFYHCTNEDDEGSCADHPSCSRSNC 231
 Db 503 CNTKINGEVMHALLRQTMHTFCFVCAACKRPFNSLFH---MEDGEPCYCKDYINLFSTKC 559
 QY 232 --C-----ARNSFMGAL-----SVLPCLLCYLPATGCVKLAQRGYDLRRPGCRCKHTN 279
 Db 560 HGCDPVEAGDKFTALGHTWHDTCYICAVCHVNLG-----OPFYSKDRPLCK-KHAH 613
 QY 280 SV 281
 Db 614 TI 615

RESULT 11

09JKS4

PRELIMINARY; PRT; 723 AA.

AC 09JKS4;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE ORACLE 1 PROTEIN (FRAGMENT).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HEART;
 RX MEDLINE=20193503; PubMed=10727866;
 RA Passier R., Richardson J.A., Olson E.N.,
 RT "Oracle, a novel PDZ-LIM domain protein expressed in heart and
 RT skeletal muscle.",
 RL Mech. Dev. 92:277-284(2000).
 CC -1- SIMILARITY: CONTAINS LIM DOMAIN(S). THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.
 DR EMBL: AF228057; AAF33847.1; -;
 DR Interpro: IPR000345; Cytc_heme_bind.
 DR Interpro: IPR003006; Ig_MHC.
 DR Interpro: IPR001781; LIM.
 DR Interpro: IPR001478; PDZ.
 DR Interpro: IPR002965; P-rich_extensn.
 DR Pfam: PF00412; LIM; 3.
 DR Pfam: PF00595; PDZ; 1.
 DR PRINTS: PRO1217; PRICHEXTENS.
 DR PRODOM: PDO00094; LIM; 3.
 DR SMART: SM00132; LIM; 3.
 DR SMART: SM00228; PDZ; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
 DR PROSITE: PS00478; LIM_DOMAIN_1; 2.
 DR PROSITE: PS50023; LIM_DOMAIN_2; 3.
 DR PROSITE: PS50106; PDZ; 1.
 KM LIM motif: Metal-binding; zinc.
 FT NON_TER 723
 SQ SEQUENCE 723 AA; 76431 MW; CC67D38AC2FFA6B6 CRC64;

Query Match

Best Local Similarity 19.5%; Score 111.5; DB 11; Length 723;
 Matches 78; Conservative 47; Mismatches 125; Indels 149; Gaps 22;

QY 1 MEPIPOSAPLTGNSV---MVQPLDLSRMSHSLQHPLTILPLDQKTSHEVNDIUNP 56
 Db 281 PAKPRPVYVTAIRPSYGVVPASTYSPSGANYSPTPTPSAPA-----YTPSP 331
 QY 57 SLATTTGPKRTGGAPELAPTPARCDDVTHHMSFGSRSSVS----- 101

```

Db 354 VENNPPQASAVSP-AAASPAASHTSYS--EGPAAPAKPRVVTASIRSVYQVPAS 410
Qy 46 -----SHENDYIDNPSLALT-----TGKRRGAGAPELAPTPA 79
Db 411 SVSPSPGANSPTPTTSPAPATYTPSPAPYTPSSPADATYTPSPAPATYTPPS 470
Qy 80 RCDODVTHHMTSFGSRPSSVS-----SSSTSSDQRLDHMAPP--PV 120
Db 471 -----AAVSGGSESASRPPWYTDSPQKFAFGKSTTVTSKOTLRGAPATYPT 520
Qy 121 ADQASP-----RAVRI-----QPKVHCOPDLKGPV-----PELDKHFLLCE 160
Db 521 GPQVTPLAGTFFQRAERFPASSRTPPLGCHCNV--IRGPFVAMGRSMHPEEPN----- 572
Qy 161 ACGKCK-----CKECASPRTPPLSCGWCNOBCLCSAOTLVN---YGTG-M 200
Db 573 -CAVCKTSLADVCFVEBQNNVYCCERCYEOFFAPICAKCNKIMGEVNHALROTWHHTTCFV 631
Qy 201 CLV-----OGIFHYCTNEDEBGCADHPCSCSRNC--C-----ARWSPGAL-----S 242
Db 632 CAACKRPFNSLFFH---MEDGEPYCEKDYINLFSTKCHGCDFFPEAGDKFIEALGHTWHD 688
Qy 243 VLPCLLCYLPATGCVKLAORGYDLRRPGCRCKHTNSV 281
Db 689 TCFICAVCHVLEGG-----QPFYSKKDKPLCK-KHAHAI 721

RESULT 12
Q9WVH2 PRELIMINARY; PRT; 723 AA.
AC Q9WVH2:
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CYPERH1.
GN ZASP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99321910; PubMed-10391924;
RA Zhou Q., Ruiz-Lozano P., Martone M.E., Chen J.;
RT "Cypher, a Striated Muscle-restricted PDZ and LIM Domain-containing
RT Protein, Binds to alpha-Actinin-2 and Protein Kinase C.";
RN J. Biol. Chem. 274:19807-19813(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Chen J., Zhou Q.;
RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS LIM DOMAIN(S). THE LIM DOMAIN BINDS 2 ZINC
CC IONS.
DR EMBL: AF114378; AAD42950.2; -.
DR HSSP: Q05158; 1QI1.
DR MGD: MGI:1344412; Zasp.
DR InterPro: IPR000345; CytC_heme_bind.
DR InterPro: IPR001781; LIM.
DR InterPro: IPR001478; PDZ.
DR InterPro: IPR002965; P_rich_extensn.
DR Pfam: PF00412; LIM; 3.
DR Pfam: PF00595; PDZ; 1.
DR PRINTS: PR01217; PRICHEXTENS.
DR ProDom: PD000094; LIM; 3.
DR SMART: SM00132; LIM; 3.
DR SMART: SM00228; PDZ; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE: PS50023; LIM_DOMAIN_2; 3.
DR PROSITE: PS50106; PDZ; 1.
DR LIM motif; Metal-binding; Zinc.
SQ SEQUENCE 723 AA; 76445 MW; CC6829FDP2FDD691 CRC64;

```

```

Query Match 6.7%; Score 111.5; DB 11; Length 723;
Best Local Similarity 19.5%; Pred. No. 0.043;
Matches 76; Conservative 47; Mismatches 125; Indels 149; Gaps 22;

Qy 1 MEPPISAPLPNSVWVOPLLDSRMSHSLQHPILTLPIQDQVT----- 45
Db 354 VENNPPQASAVSP-AAASPAASHTSYS--EGPAAPAKPRVVTASIRSVYQVPAS 410
Qy 46 -----SHENDYIDNPSLALT-----TGKRRGAGAPELAPTPA 79
Db 411 SVSPSPGANSPTPTTSPAPATYTPSPAPYTPSSPADATYTPSPAPATYTPPS 470
Qy 121 ADQASP-----RAVRI-----QPKVHCOPDLKGPV-----PELDKHFLLCE 160
Db 521 GPQVTPLAGTFFQRAERFPASSRTPPLGCHCNV--IRGPFVAMGRSMHPEEPN----- 572
Qy 161 ACGKCK-----CKECASPRTPPLSCGWCNOBCLCSAOTLVN---YGTG-M 200
Db 573 -CAVCKTSLADVCFVEBQNNVYCCERCYEOFFAPICAKCNKIMGEVNHALROTWHHTTCFV 631
Qy 201 CLV-----OGIFHYCTNEDEBGCADHPCSCSRNC--C-----ARWSPGAL-----S 242
Db 632 CAACKRPFNSLFFH---MEDGEPYCEKDYINLFSTKCHGCDFFPEAGDKFIEALGHTWHD 688
Qy 243 VLPCLLCYLPATGCVKLAORGYDLRRPGCRCKHTNSV 281
Db 689 TCFICAVCHVLEGG-----QPFYSKKDKPLCK-KHAHAI 721

RESULT 13
Q75112 PRELIMINARY; PRT; 734 AA.
AC Q75112:
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE KIAA0613 PROTEIN (FRAGMENT).
GN KIAA0613.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-BRAIN;
RX MEDLINE-98403880; PubMed-9734811;
RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. X.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RN DNA Res. 5:169-176(1998).
CC -1- SIMILARITY: CONTAINS LIM DOMAIN(S). THE LIM DOMAIN BINDS 2 ZINC
CC IONS.
DR EMBL: AB014513; BAA31588.1; -.
DR HSSP: Q05158; 1QI1.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001781; LIM.
DR InterPro: IPR001478; PDZ.
DR InterPro: IPR002965; P_rich_extensn.
DR Pfam: PF00412; LIM; 3.
DR Pfam: PF00595; PDZ; 1.
DR PRINTS: PR01217; PRICHEXTENS.
DR ProDom: PD000094; LIM; 3.
DR SMART: SM00132; LIM; 3.
DR SMART: SM00228; PDZ; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE: PS00478; LIM_DOMAIN_1; 2.

```

DR PROSITE: PS50023; LIM_DOMAIN_2; 3.
 DR PROSITE: PS50106; PDZ_1.
 KM LIM motif; Metal-binding; Zinc.
 FT NON_TER 1
 SQ SEQUENCE 734 AA; 77738 MW; 5C89AC39CC690FB8 CRC64;

Query Match 6.7%; Score 111.5; DB 4; Length 734;
 Best Local Similarity 21.3%; Pred. No. 0.044;
 Matches 77; Conservative 42; Mismatches 132; Indels 111; Gaps 19;

OY 4 PIPQAPLPNSV---WVQPLDLSRMSHRLQH---PLTILPIDQVKTSHVENDYIDNP 56
 DB 398 PAKPRPVYVTAISIRPSYGVPASTYSPGANYSPPTYPSPAPA-----YTPSP 448
 OY 57 SLALTTPKRTGAPRLATPARCDODVTHMISFGSRSSVS----- 101
 DB 449 APAYTPSPVPTYPSPAPAYTPSPAPNYPNAPSVASGGAEPASRPVWTDTSFQKFA 508
 OY 102 --SSSTSDQRLDHNAP-----PVADQASPAVRI-----QPKVHCQPLDLKG 145
 DB 509 PGASTTISIKOTLPKPGPATTPAGPOVPLARGTVQRAEFPASSRTPLCGHCNNV-IRG 567
 OY 146 PAV-----PELDKHFLLCEACGCKK-----CKEASPTLPSCWV 181
 DB 568 PFLVAMGRSMHPEEF-----TCAYCKTSLADYCFVEQNNVYGERCYEQFFAPLCAK 619
 OY 182 CNOECLCSAOTLVN---YGTG-MCLV-----QGIFHCINEDDEGSCADHPSCSRSNC 231
 DB 620 CNPKIMEVMAHLKQTHHTTCFVCAACCKRFGNSLFH--MEGEPRCEXDYINLPSTKC 676
 OY 232 --C-----AKMSFPGAL-----SVLFLCLCYLPATGCVKLAORGVDRLRRPGCRCKTN 279
 DB 677 HGGDFPEADGKFTLEALGHWHDTGFCFICAVCHWLBG-----QPTFSKRRPLCK-KHAH 730
 OY 280 SV 281
 DB 731 TI 732

RESULT 14
 O9C010 PRELIMINARY; PRT; 832 AA.
 AC O9C010;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE KIAA1683 PROTEIN (FRAGMENT).
 GN KIAA1683.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21082932; PubMed=11214970;
 RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIX.
 RT The complete sequences of 109 new cDNA clones from brain which code
 RT for large proteins in vitro."
 RL DNA Res. 7:347-355(2000).
 DR EMBL; AB051470; BAB21774.1; -.
 FT NON_TER 1
 SQ SEQUENCE 832 AA; 85845 MW; E35A1C69D5540B01 CRC64;

Query Match 6.7%; Score 111.5; DB 4; Length 832;
 Best Local Similarity 25.0%; Pred. No. 0.049;
 Matches 69; Conservative 25; Mismatches 79; Indels 103; Gaps 16;
 OY 6 POSAPLTPNSVMQPLDLSRMSHRLQH-PLT-ILPIDQVKTSHVENDYIDNPSLALTTG 63
 DB 551 PMACGVSPS--LAOPSOATGMS-SRLAHOPKTCVSLSPAOPSOAIG---VSPSLAHPSG 604

OY 64 PKTRGAPELAPTPARCDODVTHMISFSG-----RPSVSSSSS 104
 DB 605 P---HGLSPSLA-----HPYVA-SGAGSSLAAPSOQTAVSPQAQPSVASSPBG 649
 OY 105 TSSDQRLDHNAPPVADQASPAVRIQPKVHCQPLDKGPAVPELDKHFLLCEACGK 164
 DB 650 TVSSCTLASSLAEPALAPSLFP-----PSVAHTEALSL--AQSPD----- 688
 OY 165 CKCKEASPTLPSCVNCQNECLCSAOTLVNVTGTCMCLVGIFHYHCINEDDEGSCADHPC 224
 DB 689 -----TRLAPS-----RSQPSVSGTC-----HGLDEPCWGRSLDNL 721
 OY 225 SCSRNCARMSFAGALSVLFLCLCYLPATGCVKL 260
 DB 722 QCSR-----VDSVAPC-LCHPSAFYSVL 744

RESULT 15
 O9XK3 PRELIMINARY; PRT; 667 AA.
 AC O9XK3;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE VSP417-3/A-II.
 GN VSP417-3/A-II.
 OS Giardia lamblia (Giardia intestinalis).
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardiae; Giardia.
 OX NCBI_TaxID=5741;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRIS/83/HEPU/136;
 RX MEDLINE=99053029; PubMed=9836309;
 RA Ey P.L., Darby J.M., Mayrhofer G.;
 RT "Comparison of ts4417-like variant-specific surface protein (VSP)
 RT genes in Giardia intestinalis and identification of a novel locus in
 RT genetic group II isolates."
 RL Parasitology 117:0-0(0).
 DR EMBL; AF033584; AAD03497.1; -.
 DR HSP; P02468; IILE.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR002174; Furin-like.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00001; EGF-like; 1.
 DR SMART; SM00261; FU; 3.
 SQ SEQUENCE 667 AA; 69123 MW; 77C64CFE59441C0C CRC64;

Query Match 6.7%; Score 111; DB 5; Length 667;
 Best Local Similarity 22.1%; Pred. No. 0.045;
 Matches 58; Conservative 25; Mismatches 107; Indels 72; Gaps 11;

OY 6 POSAPLTPNSVMQPLDLSRMSHRLQHPLTILPIDQVKTSHVENDYIDNPSLALTTGPK 65
 DB 121 PEAALTTDSICSGDMIG-----VTIITKNAT-----YKGVDCACACTPPN 163
 OY 66 RTGAPAPLAPTPARCDOD--VTHMISFGSRSSVSSSTSSDQRL--DHMAAPPV 120
 DB 164 RITGNAGIKTATCTGDSKIKYKTDSGTSCIEFSACNSGFFSDRRKQNSDYKICPRI 223
 OY 121 ADQASPAVRIQPKVHCQPLDKGPAVPELDKHFLLCEA---CGKC----- 165
 DB 224 DDPKANTACASDN---KKPNLEGTECNSCTDHCACFAEGTCCKSSGFLDQONCV 279
 OY 166 -----KCKEASPTLPS-CMVC-----NOCCLCSAOTLVNVTGTCMCLYQGIFFH 209
 DB 280 KSDCKTEKCACTNPRANAEEVCTECVFTGHLAPTSGCVQYCGTLGNT-----YA 328
 OY 210 CTNEDEGSCADHPSCSRSNC 231
 DB 329 GTNADKNKCKE---CRYANC 346

Sat May 11 12:51:27 2002

us-10-082-902-2.rspt

Page 9

Search completed: May 8, 2002, 10:49:36
Job time: 198 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:46:43 ; Search time 9.88 Seconds

(without alignments)
1109.593 Million cell updates/sec

Title: US-10-082-902-2

Perfect score: 1657
Sequence: 1 MEPPIPQSAFLTPNSVWVQ.....SVICKASGDANTSPPDKPF 299

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	1657	100.0	299	SPY4_HUMAN	Q9C004 homo sapien
2	1554.5	93.8	300	SPY4_MOUSE	Q9P122 mus musculu
3	673.5	40.6	313	SPY2_CHICK	Q9P112 gallus gall
4	667	40.3	313	SPY1_MOUSE	Q9QXV9 mus musculu
5	641.5	38.7	315	SPY2_HUMAN	Q43597 homo sapien
6	633.5	38.2	315	SPY2_MOUSE	Q9QXV8 mus musculu
7	622.5	37.6	288	SPY3_HUMAN	Q43610 homo sapien
8	528	31.9	139	SPY1_HUMAN	Q43609 homo sapien
9	391	23.6	589	SPY1_MOUSE	Q43609 homo sapien
10	350	15.1	76	SPY1_MOUSE	Q43609 homo sapien
11	120	7.2	2476	ZAN_PIG	Q9P111 gallus gall
12	113	6.8	809	PAT3_CAEEL	Q28983 sus scrofa
13	110.5	6.7	2813	WVF_HUMAN	Q27874 caenorhabdi
14	108	6.5	2139	CRB_DROME	P04275 homo sapien
15	107.5	6.5	1700	BAR3_CHITE	P10040 drosophila
16	107	6.5	1712	TGFB_RAT	Q00376 chironomus
17	106.5	6.4	2482	WVF_PIG	Q00918 rattus norv
18	105	6.3	5376	ZAN_MOUSE	Q08833 sus scrofa
19	104	6.3	363	PSPB_CANFA	Q48879 mus musculu
20	104	6.3	618	DUL3_HUMAN	P17129 canis fami
21	103.5	6.2	1972	1972	Q9HYJ7 homo sapien
22	102.5	6.2	577	ITB6_HUMAN	Q12888 homo sapien
23	102.5	6.2	4543	LRP1_CHICK	P18563 gallus porce
24	102	6.2	1799	1799	P81577 gallus gall
25	101	6.1	2871	LMB2_MOUSE	Q61522 mus musculu
26	101	6.1	3149	TEGU_EBV	Q61554 mus musculu
27	100.5	6.1	771	ITB2_MOUSE	P13186 epstein-bar
28	100	6.0	1964	ITB4_MOUSE	P11835 mus musculu
29	99	6.0	619	MTN4_MOUSE	P11655 mus musculu
30	98	5.9	2444	MTN4_HUMAN	Q95460 homo sapien
31	97.5	5.9	614	ITB1_HUMAN	P46531 homo sapien
32	97.5	5.9	798	ITB7_HUMAN	P20313 homo sapien
33	97	5.9	1113	COR1_MOUSE	P26010 homo sapien

34	97	5.9	2700	1	ZAN_HUMAN	Q9Y493 homo sapien
35	96.5	5.8	2569	1	LMA3_MOUSE	Q61789 mus musculu
36	96.5	5.8	2871	1	FBN1_HUMAN	P35555 homo sapien
37	96.5	5.8	3110	1	LMA2_HUMAN	P24043 homo sapien
38	96	5.8	457	1	ODR7_CAEEL	P41933 caenorhabdi
39	96	5.8	965	1	YNC3_TEAST	P53971 saccharomyc
40	96	5.8	2813	1	WVF_CANFA	Q28295 canis fami
41	95.5	5.8	1776	1	POLR_OYMY	P20127 onchis yell
42	95.5	5.8	2531	1	MTN1_RAT	Q07008 rattus norv
43	95	5.7	589	1	DUL3_RAT	Q88671 rattus norv
44	95	5.7	769	1	ITB2_BOVIN	P32592 bos taurus
45	95	5.7	769	1	ITB2_HUMAN	P05107 homo sapien

ALIGNMENTS

RESULT 1
SPY4_HUMAN STANDARD; PRT; 299 AA.

AC Q9C004; Q9C003;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE SPROUTY HOMOLOG 4 (SPRY-4).
GN SPY4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid-9606;

RA [1]
RA Leeksa O.C., van Achterberg T.A.E., Spaargaren M.,
RA van den Borne A.E.G., Pannekoek H., de Vries C.J.M.;
RA "Identification of a novel human Sprouty homolog";
RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: MAY FUNCTION AS AN ANTAGONIST OF FIBROBLAST GROWTH
CC FACTOR (FGF) PATHWAYS AND MAY NEGATIVELY MODULATE RESPIRATORY
CC ORGANOREGULATION (BY SIMILARITY).
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- DOMAIN: THE CYS-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF
CC THE PROTEIN TO THE MEMBRANE RUPTURES.
CC -1- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: AF227516; AAK0652.1; ALT-INIT.
CC EMBL: AF227517; AAK0653.1; -
CC Developmental protein; Membrane; Alternative splicing.
CC DOMAIN 97 107
CC POLY-SER.
CC FT DOMAIN 159 283
CC CYS-RICH
CC FT VARSPIC 97 106
CC MISSING (IN ISOFORM C).
CC FT VARSPIC 107 299
CC MISSING (IN ISOFORM C).
CC SQ SEQUENCE 299 AA; 32541 MW; 105F6F1BE9F7B6C3 CRC64;

Query Match 100.0%; Score 1657; DB 1; Length 299;
Best Local Similarity 100.0%; Pred. No. 5.9e-118;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPPIPQSAFLTPNSVWVQPLDLSRMSHRLQHPPLTLPIDQYKTSVENDYIDNPSLAL 60
QY |||||
Db 1 MEPPIPQSAFLTPNSVWVQPLDLSRMSHRLQHPPLTLPIDQYKTSVENDYIDNPSLAL 60
QY 61 TTGPKRTGCAFLAPLPACDDDDYTHHMISFGGRSSVSSSSSTSSDORLIDHMAAPPY 120

```

Db 61 TTGPKRTGGAPLAPTPARCDQVTHHMFISGRPSVSSSSSTSDQRLDHPAPPV 120
Cc 121 ADQASPRAVRIRQPKVHCOPDLKGPVAPPDLKHFLLEACGCKCKEASPTLPSCW 180
Cc 121 ADQASPRAVRIRQPKVHCOPDLKGPVAPPDLKHFLLEACGCKCKEASPTLPSCW 180
Db 121 ADQASPRAVRIRQPKVHCOPDLKGPVAPPDLKHFLLEACGCKCKEASPTLPSCW 180
Cc 181 VCNQECLESAOTLVNVTGCMCLVOCIFHCTNEDEGSCADHPSCSRSNCCARMSFMA 240
Cc 181 VCNQECLESAOTLVNVTGCMCLVOCIFHCTNEDEGSCADHPSCSRSNCCARMSFMA 240
Db 181 VCNQECLESAOTLVNVTGCMCLVOCIFHCTNEDEGSCADHPSCSRSNCCARMSFMA 240
Cc 241 LSVVLPCLLCYLPATGCVKLAQRGYDRLRRPGRCKHTNSVICKAASGADAKTSRDPKF 299
Cc 241 LSVVLPCLLCYLPATGCVKLAQRGYDRLRRPGRCKHTNSVICKAASGADAKTSRDPKF 299
Db 241 LSVVLPCLLCYLPATGCVKLAQRGYDRLRRPGRCKHTNSVICKAASGADAKTSRDPKF 299

RESULT 2
SPY2_MOUSE STANDARD; PRT; 300 AA.
AC Q9WTP2; Q9QXV7;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SPROUTY HOMOLOG 4 (SPRY-4).
GN SPRY4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.
NCBI_TaxID=10090;
RN [1]
RX MEDLINE=99264295; PubMed=10330503;
RA de Maximy A.A., Nakatake Y., Moncada S., Itoh N., Thierly J.P.,
RA Bellusci S.;
RT "Cloning and expression pattern of a mouse homologue of drosophila
RT sprouty in the mouse embryo."
RL Mech. Dev. 81:213-216(1999).
RN [2]
RX SEQUENCE FROM N.A.
RX MEDLINE=99429807; PubMed=10498682;
RA Minowada G., Jarvis L.A., Chi C.L., Neubuser A., Sun X., Hachon N.,
RA Krasnow M.A., Martin G.R.;
RT "Vertebrate sprouty genes are induced by FGF signaling and can cause
RT chondrodysplasia when overexpressed."
RL Development 126:4465-4475(1999).
Cc -1- FUNCTION: MAY FUNCTION AS AN ANTAGONIST OF FIBROBLAST GROWTH
Cc FACTOR (FGF) PATHWAYS AND MAY NEGATIVELY MODULATE RESPIRATORY
Cc ORGANOGENESIS.
Cc -1- SUBCELLULAR LOCATION: FOUND IN THE CYTOPLASM IN UNSTIMULATED CELLS
Cc BUT IS TRANSLOCATED TO THE MEMBRANE RUFFLES IN CELLS STIMULATED
Cc WITH EGF (EPIDERMAL GROWTH FACTOR).
Cc -1- TISSUE SPECIFICITY: EXPRESSED IN THE EMBRYO AND ADULT TISSUES
Cc INCLUDING HEART, BRAIN, LUNG, KIDNEY, AND SKELETAL MUSCLE.
Cc -1- DEVELOPMENTAL STAGE: AT 8 DPC EXPRESSED IN THE LATERAL PLATE
Cc MESODERM OF THE PRIMITIVE STREAK. AT 9.5 AND 10.5 DPC EXPRESSED IN
Cc THE NASAL PLACODES, MAXILLARY AND MANDIBULAR PROCESSES, POSTERIOR
Cc PART OF THE HYOID ARCH AND THE PROGRESS ZONE OF THE LIMB BUDS AND
Cc THE PRESOMITIC MESODERM. AT 11.5 DPC EXPRESSED IN THE DORSO-
Cc LATERAL REGION OF THE SOMITES (MOSTLY IN THE MYOTOME) AND IN THE
Cc OTIC VESICLE. AT 11.5 AND 12.5 DPC EXPRESSED IN THE DISTAL LUNG
Cc MESENCHYME, WITH A STRONG EXPRESSION IN THE ACCESSORY LOBE OF THE
Cc LUNG.
Cc -1- INDUCTION: BY FGF SIGNALING.
Cc -1- DOMAIN: THE CYS-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF
Cc THE PROTEIN TO THE MEMBRANE RUFFLES.
Cc -1- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.
Cc -----
Cc THIS SWISS-PROT entry is copyright. It is produced through a collaboration
Cc between the Swiss Institute of Bioinformatics and the EMBL outstation -
Cc the European Bioinformatics Institute. There are no restrictions on its
Cc use by non-profit institutions as long as its content is in no way
Cc modified and this statement is not removed. Usage by and for commercial
Cc entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```

```

Cc or send an email to license@isb-sib.ch.
Cc -----
Cc EMBL: AB019280; BAA77689.1; -
Cc DR EMBL: AF176906; AAD56007.1; -
Cc DR MGI: 1345144; Spry4.
Cc KW Developmental protein; Membrane.
Cc FT DOMAIN 98 108 POLY-SER.
Cc FT DOMAIN 160 284 CYS-RICH.
Cc FT CONFLICT 10 10 V -> F (IN REF. 2).
Cc SQ SEQUENCE 300 AA; 32523 MW; DA963036EFC0E73F CRC64;

Query Match 93.8%; Score 1554.5; DB 1; Length 300;
Best Local Similarity 92.7%; Pred. No. 36-110;
Matches 278; Conservative 9; Mismatches 12; Indels 1; Gaps 1;

Cc 1 MEPPITQSA-PLTPNSVMQPLDLSRMSIRLOHPLTLPIDQVKTSHENDYIDNPSLA 59
Cc 1 MEPPITQSSVVPVSSVMQPLDLSRAPHSRLQHPLTLPIDQKTSHENDYIDNPSLA 60
Db 1 MEPPITQSSVVPVSSVMQPLDLSRAPHSRLQHPLTLPIDQKTSHENDYIDNPSLA 60
Cc 60 LTTGPKRTGGAPLAPTPARCDQVTHHMFISGRPSVSSSSSTSDQRLDHPAPPV 119
Cc 61 PATGPKRPGGAPLAPTPARCDQVTHHMFISGRPSVSSSSSTSDQRLDHPAPPV 120
Db 61 PATGPKRPGGAPLAPTPARCDQVTHHMFISGRPSVSSSSSTSDQRLDHPAPPV 120
Cc 120 VADQASPRAVRIRQPKVHCOPDLKGPVAPPDLKHFLLEACGCKCKEASPTLPSC 179
Db 120 VADQASPRAVRIRQPKVHCOPDLKGPVAPPDLKHFLLEACGCKCKEASPTLPSC 180
Cc 121 VADQASPRAVRIRQPKVHCOPDLKGPVAPPDLKHFLLEACGCKCKEASPTLPSC 180
Cc 180 VCNQECLESAOTLVNVTGCMCLVOCIFHCTNEDEGSCADHPSCSRSNCCARMSFMA 239
Cc 181 VCNQECLESAOTLVNVTGCMCLVOCIFHCTNEDEGSCADHPSCSRSNCCARMSFMA 240
Cc 240 ALSVLPCLLCYLPATGCVKLAQRGYDRLRRPGRCKHTNSVICKAASGADAKTSRDPKF 299
Cc 241 ALSVLPCLLCYLPATGCVKLAQRGYDRLRRPGRCKHTNSVICKAASGADAKTSRDPKF 300
Db 241 ALSVLPCLLCYLPATGCVKLAQRGYDRLRRPGRCKHTNSVICKAASGADAKTSRDPKF 300

RESULT 3
SPY2_CHICK STANDARD; PRT; 313 AA.
AC Q9PTL2;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SPROUTY HOMOLOG 2 (SPRY-2).
GN SPRY2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE=99429807; PubMed=10498682;
RA Minowada G., Jarvis L.A., Chi C.L., Neubuser A., Sun X., Hachon N.,
RA Krasnow M.A., Martin G.R.;
RT "Vertebrate sprouty genes are induced by FGF signaling and can cause
RT chondrodysplasia when overexpressed."
RL Development 126:4465-4475(1999).
RN [2]
RX SEQUENCE FROM N.A.
RX TISSUE=Embryo;
RX MEDLINE=20130796; PubMed=10662503;
RA Chambers D., Medhurst A.D., Walsh F.S., Price J., Mason I.;
RT "Differential display of genes expressed at the midbrain-hindbrain
RT junction identifies sprouty2: an Fgf8-inducible member of a family of
RT intracellular GGF antagonists."
RL Mol. Cell. Neurosci. 15:22-35(2000).
Cc -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN FGF-MEDIATED PATTERNING OF
Cc THE MID/HINDRAIN REGION BY ACTING TO MODULATE THE SIGNALING
Cc EFFECTS OF FGF8 THROUGH PARTICIPATION IN A REGULATORY NEGATIVE
Cc FEEDBACK LOOP.
Cc -1- SUBCELLULAR LOCATION: FOUND IN THE CYTOPLASM IN UNSTIMULATED CELLS

```

CC BUT IS TRANSLOCATED TO THE MEMBRANE RUFLIES IN CELLS STIMULATED
 CC WITH EGF (EPIDERMAL GROWTH FACTOR).
 CC -1- TISSUE SPECIFICITY: BRAIN AND INTERLIMB REGION.
 CC -1- DEVELOPMENTAL STAGE: AT THE 4-TO 5-SOMITE STAGE (4/5S) FOUND IN
 CC THE EMBRYO IN SCATTERED CELLS ACROSS THE NEURAL PLATE IN THE
 CC PRESUMPTIVE MID/ HINDERBRAIN REGION. AT 7/8S FOUND IN THE ISTHMUS
 CC AND THROUGHOUT THE PRESUMPTIVE R1 TERRITORY. BETWEEN 10-14S STAGE
 CC FOUND THROUGHOUT THE R1 REGION AND AT THE ISTHMIC CONSTRICTION. BY
 CC 26S THE ANTERIOR LIMIT OF EXPRESSION EXTENDS INTO THE POSTERIOR
 CC MIDBRAIN REGION AND THIS PATTERN OF EXPRESSION IS MAINTAINED AT
 CC LATER STAGES.
 CC -1- INDUCTION: BY EGF SIGNALING.
 CC -1- DOMAIN: THE CYS-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF
 CC THE PROTEIN TO THE MEMBRANE RUFLIES.
 CC -1- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF176904; AAD56005.1; -
 CC Developmental protein; Membrane.
 CC FT DOMAIN 176 299 CYS-RICH.
 CC SQ SEQUENCE 313 AA; 34417 MW; 614C335ED0C36E83 CRC64;

Query Match 40.6%; Score 673.5; DB 1; Length 313;
 Best Local Similarity 44.8%; Pred. No. 5.7e-44;
 Matches 138; Conservative 39; Mismatches 76; Indels 55; Gaps 9;

QY 19 OPTLSRMSHSR-----LQHPLTLPIDOVKTSHEVDYIDNPDLATLTGKRRR 68
 DB 12 QALLQRRSGRPHGEPDLRDVLTQVHILSDQIRAINKNTVEYEGPVAPRGK--- 68
 QY 69 GGAPELAPTP-----ARCDDVTH-----HWISFGSPSSVS 103
 DB 69 -SAPRSASQPKSERPHGLDEHRHFGVHQTGHASPRAPLRSISVSGSSSTRTS 127
 QY 104 STSSDRLDHPVADQASPRVRIQPKVYHCOPDLKGPVPP-----ELDKHFLLC 159
 DB 128 SNSEQRLLG-SSSGPVAD-----GIYRQPK-----SELKSELPLSKEDLGASYSRC 176
 QY 160 ECGKCKECCASPRTPSCWVNOECISAOITLVNYGCMCYOGIFHCHNDEDESGC 219
 DB 177 EDGKCKECCASPRTPSCWVNOECISAOITLVNYGCMCYOGIFHCHNDEDESGC 235
 QY 220 ADHPCSSRSCARFSGALSVLPCLCYLPATGCVKLAORGYDRLRRPCRCRKHNN 279
 DB 236 ADHPCSSRSCARFSGALSVLPCLCYLPATGCVKLAORGYDRLRRPCRCRKHNN 295
 QY 280 SVICKAAS 287
 DB 296 TVCKKVP 303
 RESULT 4
 SPYL_MOUSE STANDARD; PRT; 313 AA.
 AC 090XV9;
 DT 20-AUG-2001 (Rel. 40; Created)
 DT 20-AUG-2001 (Rel. 40; Last sequence update)
 DT 20-AUG-2001 (Rel. 40; Last annotation update)
 DE SPROUTY HOMOLOG 1 (SPRT-1).
 GN SPRT1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
 RA MEDLINE-99429807; PubMed-10498682;
 RA Minowada G., Jarvis L.A., Chl C.L., Neubuser A., Sun X., Hachon N.,
 RA Kirsnow M.A., Martin G.R.;
 RT "Vertebrate sprouty genes are induced by FGF signaling and can cause
 RT chondrodysplasia when overexpressed."
 RL Development 126:4465-4475(1999).
 CC -1- FUNCTION: MAY FUNCTION AS AN ANTAGONIST OF FIBROBLAST GROWTH
 CC FACTOR (FGF) PATHWAYS AND MAY NEGATIVELY MODULATE RESPIRATORY
 CC ORGANORESIS.
 CC -1- SUBCELLULAR LOCATION: FOUND IN THE CYTOPLASM IN UNSTIMULATED CELLS
 CC BUT IS TRANSLOCATED TO THE MEMBRANE RUFLIES IN CELLS STIMULATED
 CC WITH EGF (EPIDERMAL GROWTH FACTOR).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE EMBRYO AND ADULT TISSUES
 CC INCLUDING HEART, BRAIN, LUNG, KIDNEY, AND SKELETAL MUSCLE.
 CC -1- DEVELOPMENTAL STAGE: AT E8.5 EXPRESSED IN THE PRIMITIVE STREAK,
 CC ROSTRAL FOREBRAIN, CELLS LATERAL TO THE POSTERIOR HINDERBRAIN,
 CC ANTERIOR HINDERBRAIN AND DEVELOPING MIDBRAIN. AT E9.5 CONTINUES TO
 CC BE EXPRESSED IN THE ROSTRAL FOREBRAIN AND PRIMITIVE STREAK, AND IS
 CC ALSO DETECTED IN THE BRANCHIAL ARCHES AND THE FORELIMB BUD. AT
 CC E10.5 EXPRESSED IN THE SOMITES, FRONTONASAL PROCESSES, TAILBUD,
 CC AND HINDLIMB BUD.
 CC -1- INDUCTION: BY EGF SIGNALING.
 CC -1- DOMAIN: THE CYS-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF
 CC THE PROTEIN TO THE MEMBRANE RUFLIES.
 CC -1- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF176903; AAD56004.1; -
 CC MGD: MGI:1345139; Spry1.
 CC Developmental protein; Membrane.
 CC FT DOMAIN 116 125 POLY-SER.
 CC SQ SEQUENCE 313 AA; 34004 MW; CD3C2F7022C0B99 CRC64;

Query Match 40.3%; Score 667; DB 1; Length 313;
 Best Local Similarity 44.6%; Pred. No. 1.8e-43;
 Matches 135; Conservative 46; Mismatches 78; Indels 44; Gaps 8;

QY 16 VAVQPL---DSRMSHRLQHPITLPIIDVKTSHVENDYIDNPDLATLTGKRRGAP 72
 DB 14 VVIQPAVAGRGRLDYDRTPATILSLDIKAIKINSNEYTEGSPVA-----RRAP 65
 QY 73 ELAPPARDDOVTHHWI-----SFGSRP-----SSVSSSST 105
 DB 66 RTAPRPEK--QRTHEIIPANVNSSEYHRPASHPCNANGSVLSRSTSGSAASSSSSSSV 123
 QY 106 SSDRLDHPMAP-IVADQASPRVRIQPKVYHCOPDLKGPVPPDLKHFLLCEACGK 164
 DB 124 SSEGGLGSRSPTRIPGRHSRDRVIRIQKOLIVE--DLKA-SLKEDPTQHKFCEQCK 180
 QY 165 CKKRCASPRTPISGWCNOCGLCSAOTLVNCTCMCLYOGIFHYCHTNEDEGSCADHPC 224
 DB 181 CKGRCETAPRALPSCLACDRCLCSAESVVEGTGCMCLYKGFYHCSNDDDGSGYSIDPC 240
 QY 225 SCRSRSCARFSGALSVLPCLCYLPATGCVKLAORGYDRLRRPCRCRKHNN 284
 DB 241 SCOSHCSCRYLCKGALSLCLPCLCYLPAPKCALCKRCQCYDWTIRPCRCRNNSTVYCK 300
 QY 285 AAS 287
 DB 301 LES 303
 RESULT 5

ID	SPY2_HUMAN	STANDARD:	PRT:	315 AA.
AC	043597:			
DT	20-AUG-2001 (rel. 40, Created)			
DT	20-AUG-2001 (rel. 40, Last sequence update)			
DT	20-AUG-2001 (rel. 40, Last annotation update)			
DE	SPROUTY HOMOLOG 2 (SPRY-2).			
OS	SPRY2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxId=9606;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RA	Macdonen-98117253; PubMed=9458049;			
RA	Hachon N., Kramer S., Sutherland D., Hiromi Y., Krasow M.A.;			
RT	*Sprouty encodes a novel antagonist of EGF signalling that patterns			
RT	apical branching of the Drosophila airways.*;			
RL	Cell 92:253-263(1998).			
RN	[2]			
RP	FUNCTION, SUBCELLULAR LOCATION, AND DOMAIN.			
RA	Medline=20490720; PubMed=10887178;			
RA	Lim J., Mong E.S.M., Ong S.H., Yusoff P., Low B.C., Guy G.R.;			
RT	*Sprouty proteins are targeted to membrane ruffles upon growth factor			
RT	receptor tyrosine kinase activation. Identification of a novel			
RT	translocation domain.*;			
RL	J. Biol. Chem. 275:32837-32845(2000).			
CC	-1- FUNCTION: MAY FUNCTION AS AN ANTAGONIST OF FIBROBLAST GROWTH			
CC	FACTOR (FGF) PATHWAYS AND MAY NEGATIVELY MODULATE RESPIRATORY			
CC	ORANOGENESIS.			
CC	-1- SUBCELLULAR LOCATION: ASSOCIATED WITH MICROTUBULES IN UNSTIMULATED			
CC	CELLS BUT IS TRANSLOCATED TO THE MEMBRANE RUFFLES IN CELLS			
CC	STIMULATED WITH EGF (EPIDERMAL GROWTH FACTOR).			
CC	-1- INDUCTION: BY EGF SIGNALING.			
CC	-1- DOMAIN: THE CYS-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF			
CC	THE PROTEIN TO THE MEMBRANE RUFFLES.			
CC	-1- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/).			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL: AF039843; AAC04258.1; -.			
DR	MIM: 602456; -.			
KW	Developmental protein; Membrane; Microtubules.			
FT	DOMAIN 125 131 POLY-SER.			
FT	DOMAIN 178 301 CYS-RICH.			
SO	SEQUENCE 315 AA; 34688 MW; 8CC6256929D91A7E CRC64;			
Query Match 38.7%; Score 641.5; DB 1; Length 315;				
Best Local Similarity 43.6%; Pred. NO. 1.5e-41;				
Matches 130; Conservative 45; Mismatches 84; Indels 39; Gaps 8;				
OY	19	QPLDLSRMSHNR-----LQNPRLTLPIDQYKTSIVENDYIDNPSLALTGPK---	65	
DB	12	QPLQTPRDGGGQNGEPDRDALTQGVHLSDLDRAIRLNTMEYEGPVRPRPKAP	71	
OY	66	-----RTRGAPDELAPTP-----ARCQDVTHMHSISGRSSVSSSSSS	107	
DB	72	RPTQHKHEHRLGRLGPEHQPRRLQSHQVNSAKRALSLSSISIVSSGSSSTRTTSSSSS	131	
OY	108	DQRLI-DHMAPRVADQASPAVRVITQPKVNHQQLRLGPAVRPELDHFLLEACGRCK	166	
DB	132	EQRLLGSSFSFGSPVAD----GIIRVQPK-SELKPELKL-PLSKEDLDGLHAYVCEDCGRCK	185	
OY	167	CKEASPTPLTSCVNCWNCQECISAGTLVNVNGTCKMLVQGIYHCTNEDDEGSCADHPSC	226	
DB	186	CKEYTPRLPSDMLCDKQCLSAQNVNIDYGTGVCCVGLGELHGSN-DDEDNCADNPSC	244	

QY	227	SNSNCCANSPFGALSYVPLCLCYCPAPGCVKLAQRGVDRILRRPCCRCKHTNTSVYC	284
DB	245	SOSHCTRMSANGVSNLFLPCIMCTIPLANGCLKLCOGCIDRVNRPGCRCKNSMTVCCK	302
RESULT_6			
ID	SPY2_MOUSE	STANDARD:	PRT: 315 AA.
AC	09OXV8: 09MU09;		
DT	20-AUG-2001 (rel. 40, Created)		
DT	20-AUG-2001 (rel. 40, Last sequence update)		
DE	20-AUG-2001 (rel. 40, Last annotation update)		
GN	SPROUTY HOMOLOG 2 (SPRY-2).		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
OX	NCBT_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99429807; PubMed=10498682;		
RA	Minowada G., Jarvis L.A., Chi C.L., Neubueser A., Sun X., Haeohen N.,		
RA	Krisnow M.A., Martin G.R.;		
RT	"Vertebrate sprouty genes are induced by FGF signalling and can cause		
RT	chondrodysplasia when overexpressed.";		
RL	Development 126:4465-4475(1999).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=SWISS WEBSTER;		
RX	Tefft J.D., Lee M., Smith S., Leinwand M., Zhao J., Bringas P. Jr.,		
RA	Crowe D.L., Warburton D.;		
RT	"Conserved function of mSpry-2, a murine homolog of Drosophila		
RT	sprouty, which negatively modulates respiratory organogenesis.";		
RL	Curr. Biol. 9:219-222(1999).		
CC	-!- FUNCTION: MAY FUNCTION AS AN ANTAGONIST OF FIBROBLAST GROWTH		
CC	FACTOR (FGF) PATHWAYS AND MAY NEGATIVELY MODULATE RESPIRATORY		
CC	ORGANOGENESIS.		
CC	-!- SUBCELLULAR LOCATION: ASSOCIATED WITH MICROTUBULES IN UNSTIMULATED		
CC	CELLS BUT IS TRANSLOCATED TO THE MEMBRANE RUPTURES IN CELLS		
CC	STIMULATED ITH EGF (EPIDERMAL GROWTH FACTOR).		
CC	-!- TISSUE SPECIFICITY: IN ADULT, HIGHLY EXPRESSED IN BRAIN, LUNG,		
CC	HEART AND AT LOWER LEVELS IN SKELETAL MUSCLE AND KIDNEY. IN		
CC	EMBRYO, HIGHLY EXPRESSED IN LUNG EPITHELIAL CELLS, PRIMARILY IN		
CC	THE DISTAL AIRWAYS.		
CC	-!- DEVELOPMENTAL STAGE: AT E8.5 EXPRESSED IN THE PRIMITIVE STREAK,		
CC	POSTAL FOREBRAIN, CELLS LATERAL TO THE POSTERIOR HINDBRAIN,		
CC	ANTERIOR HINDBRAIN AND DEVELOPING MIDBRAIN. AT E9.5 CONTINUES TO		
CC	BE DETECTED IN THE ROSTRAL FOREBRAIN AND PRIMITIVE STREAK, AND IS		
CC	ALSO DETECTED IN THE BRANCHIAL ARCHES AND THE FORLUMB BUD. AT		
CC	E10.5 EXPRESSED IN THE SOMITES, FRONTONASAL PROCESSES, TAILBUD,		
CC	AND HINDLIMB BUD.		
CC	-!- INDUCTION: BY EGF SIGNALING.		
CC	-!- DOMAIN: THE CY5-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF		
CC	THE PROTEIN TO THE MEMBRANE RUPTLES.		
CC	-!- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.		
CC	-----		
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation --		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; AF176905; AAD56006.1; -		
DR	EMBL; AF153084; AAD34167.1; -		
DR	MCD; MG1:1345138; Spry2.		
KW	Developmental protein; Membrane; Microtubules.		
FT	DOMAIN 124 130		
FT	DOMAIN 178 301		
FT	CONFLICT 163 163		
FT	CONFICT V -> I (IN REF. 2).		

50 SEQUENCE 315 AA; 34623 MW; 81514698B809A7 CRC64;

Query Match
 Best Local Similarity 43.3%; Score 633.5; DB 1; Length 315;
 Matches 129; Conservative 46; Mismatches 84; Indels 39; Gaps 9;

19 QPLDNRSHSR-----LQHPILPIDQVTSVENDYIDNSLATTGPK-- 65
 12 QPLLOTADHSGRORGEPPDRDLTQOVHVLSDQIRAIRNTEGPTVPRGLKAP 71
 66 -----FRG-----GAPELAPT---PARCDDYTHHMTSGRSSSSSTSD 108
 72 RPTQHKERLHGLPEHROPRLDPSQVHSSRAPLSKSTVSSGSRSTRTSSSSSE 131
 109 QRL--DHMAPPVADQSPRAVRITQPKVHCOPLDLKGAPELDHFLKCEACGCK 166
 132 QRLGSPFSHGPAAAD-----GIIRQPK-SELKPGDYK-PLSKDGLGHAIRCEDCGCK 185
 167 CKECASPTLTSQVNCNOELCSAOTLVNCTCMQVGIYHCTNEDEGSCADHPSC 226
 186 KCECTPPRLPSDMDQCCSASQNVYDGTCCVCCVGLFHCSSN--DDEDCADNPSC 244

227 SRNCCARMSFGALSVLPCLCYLPATGCVKLAQREGDRLRPGCRCKHTNVICK 284
 245 SQSHCTMSAMGVSLFPLCLNCTLPKAGCLKLCQGCYDVRNRPCKCKNSNTVCK 302

RESULT 7
 SPY3_HUMAN STANDARD; PRT; 288 AA.
 AC 043610;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DE 20-AUG-2001 (Rel. 40, Last annotation update)
 GN SPROUTY HOMOLOG 3 (SPRY-3).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ciccocioppa A., D'Esposito M., Esposito T., Gianfrancesco F.,
 RA Migliaccio C., Miano M.G., Matarazzo M.R., Vacca M., Franz A.,
 RA Cucchiarese M., Cocchia M., Curci A., Terraciano A., Torino A.,
 RA Cocchia S., Mercadante G., Pannone E., Archidiacono N., Rocchi M.,
 RA Schlesinger D., D'Urso M.;
 RA "Differentially regulated and evolved genes in the fully sequenced
 RT Xq/Yq pseudautosomal region";
 RT Hum. Mol. Genet. 9:395-401(2000).
 RL [2]
 RN SEQUENCE OF 92-191 FROM N.A.
 RP MEDLINE-98117253; PubMed-9458049;
 RA Hachohen N., Kramer S., Sutherland D., Hironi Y., Krasnow M.A.;
 RA "Sprouty encodes a novel antagonist of FGF signaling that patterns
 RT apical branching of the Drosophila airways";
 RT Cell 92:253-263(1998).
 RL -1- FUNCTION: MAY FUNCTION AS AN ANTAGONIST OF FIBROBLAST GROWTH
 CC FACTOR (FGF) PATHWAYS AND MAY NEGATIVELY MODULATE RESPIRATORY
 CC ORGANOGENESIS.
 CC -1- SUBCELLULAR LOCATION: FOUND IN THE CYTOPLASM IN UNSTIMULATED CELLS
 CC BUT IS TRANSLOCATED TO THE MEMBRANE RUFLIES IN CELLS STIMULATED
 CC WITH EGF (EPIDERMAL GROWTH FACTOR).
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED; PARTICULARLY IN THE FETAL
 CC TISSUES.
 CC -1- INDUCTION: BY FGF SIGNALING.
 CC -1- DOMAIN: THE CYS-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF
 CC THE PROTEIN TO THE MEMBRANE RUFLIES.
 CC -1- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AJ271735; CAB96768.1; -
 DR EMBL; AF041038; AAC39567.1; -
 DR MIM; 602467; -
 DR InterPro: IPR003019; Metallothion.
 DR Pfam: PF00131; metalthio; 1.
 KW Developmental protein; Membrane.
 FT DOMAIN 147 270
 FT CONFLICT 92 95 TRAS -> PLPL (IN REF. 2).
 50 SEQUENCE 288 AA; 31221 MW; 62C39F5E897D99B8 CRC64;

Query Match
 Best Local Similarity 41.6%; Score 622.5; DB 1; Length 288;
 Matches 123; Conservative 40; Mismatches 62; Indels 71; Gaps 9;

37 ILPIDQVTSVENDYIDNP-----SLA-LTPGKRTRGAP 73
 12 ILPLEQLRSTHASNDRVERPPAPCKQALSSPSLIVQTHKSDWSLATWPLPSLQCHQ 71
 74 LAPTPARCDQDVTNHHMTSFGSRPSSVSS--SSTSSQRLDLHMAPPVADQSPRAVR 130
 72 LQPLPQLHLSQ-----SSIASMSHSTTADQRLASITPSP-----SGQSIIR 114
 131 IQPKVHVCQPLDLKGAVPPELD-----KHFLLCEACGCKCKEBCASPT 175
 115 TOP-----GAGVHPRKADGALKGEAQSAGHPSEHLFICEEGRCCKVPTAAR 163
 176 LPSCWVNOELCSAOTLVNCTCMQVGIYHCTNEDEGSCADHPGSCSNCCARW 235
 164 LPSCWLNQRCLESABSLDGTCLCCVKGFLYHCSY--DDEDCADBPCCGSGPSCFVR 222

236 SFGALSVLPCLCYLPATGCVKLAQREGDRLRPGCRCK-HTNSVYCKAASGDA 290
 223 AAMSLISLPLPCLCCVLPTRGCHLHCQGYDLSLRPGCRCKRHTNVCRISSGSA 278

RESULT 8
 SPY1_HUMAN STANDARD; PRT; 139 AA.
 ID SPY1_HUMAN
 AC 043609;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DE 20-AUG-2001 (Rel. 40, Last annotation update)
 GN SPROUTY HOMOLOG 1 (SPRY-1) (FRAGMENT).
 DE SPROUTY.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hachohen N., Kramer S., Sutherland D., Hironi Y., Krasnow M.A.;
 RA "Sprouty encodes a novel antagonist of FGF signaling that patterns
 RT apical branching of the Drosophila airways";
 RT Cell 92:253-263(1998).
 RL -1- FUNCTION: MAY FUNCTION AS AN ANTAGONIST OF FIBROBLAST GROWTH
 CC FACTOR (FGF) PATHWAYS AND MAY NEGATIVELY MODULATE RESPIRATORY
 CC ORGANOGENESIS.
 CC -1- SUBCELLULAR LOCATION: FOUND IN THE CYTOPLASM IN UNSTIMULATED CELLS
 CC BUT IS TRANSLOCATED TO THE MEMBRANE RUFLIES IN CELLS STIMULATED
 CC WITH EGF (EPIDERMAL GROWTH FACTOR).
 CC -1- INDUCTION: BY FGF SIGNALING.
 CC -1- DOMAIN: THE CYS-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF
 CC THE PROTEIN TO THE MEMBRANE RUFLIES.
 CC -1- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF041037; AAC39566.1; -
DR MIM: 602465; -
KM Developmental protein; Membrane.
FT NON_TER 1 130 CYS-RICH.
SQ DOMAIN 1 130 CYS-RICH.
SQ SEQUENCE 139 AA; 15359 MW; BECDAAE42B322220 CRC64;

Query Match 31.9%; Score 528; DB 1; Length 139;
Best Local Similarity 62.6%; Pred. No. 2,3e-33;
Matches 87; Conservative 24; Mismatches 24; Indels 4; Gaps 1;

QY 159 CEAGKCKCKEASPRTPSCWVNCNOCISAGTIVNYGTCMCTVGGIFVHCHNEDDGS 218
DB 1 CEAGKCKCKEASPRTPSCWVNCNOCISAGTIVNYGTCMCTVGGIFVHCHNEDDGS 60
QY 219 CADHPCSCSRNCCARMSFGALSVVLPCLLCYLPATGCVKLAQRGRLRPRGCRCKHT 278
DB 61 YSDNPGSCSGSHCCSRILCKGMSLPLPCLLCYLPATGCVKLAQRGRLRPRGCRCKHT 120
QY 279 NSVYCKAAS---GDAKTS 293
DB 121 NTYVCKLESCPSRGGKPS 139

RESULT 9
SPY_DROME STANDARD; PRT; 589 AA.
AC 044783; Q9VZP7;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROTEIN SPROUTY (SPRY).
GN SPRY OR CG1921.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RX STRAIN=CANTON-S.
RP SEQUENCE FROM N.A.
RC MEDLINE=98117253; PubMed=9458049;
RA Hecohen N., Kramer S., Sutherland D., Hiromi Y., Krasnow M.A.;
RT "Sprouty encodes a novel antagonist of FGF signaling that patterns
RT apical branching of the Drosophila airways."
RL Cell 92:253-263(1998).
RN 121
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anagnostou C., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Morten J.R., Yandell M.D., Zhang O., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agayari A., An H.-J., Andrews-Plamkoc C., Baldwin D.,
RA Ballow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Butts R.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Dey A.D., Dew I., Dietz S.M.,

```

```

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster G., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegyan C.,
RA Jaimel B.E., Kalush F., Karpén G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jaimel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mervulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nussken D.R., Pacled J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN 131
RP FUNCTION.
RX MEDLINE=99387981; PubMed=10457022;
RA Reich A., Sapir A., Shilo B.-Z.;
RT "Sprouty is a general inhibitor of receptor tyrosine kinase
RT signaling."
RL Development 126:4139-4147(1999).
RN 141
RP FUNCTION.
RX MEDLINE=99244704; PubMed=10226010;
RA Kramer S., Okabe M., Hecohen N., Krasnow M.A., Hiromi Y.;
RT "Sprouty, a common antagonist of FGF and EGF signaling pathways in
RT Drosophila."
RL Development 126:2515-2525(1999).
RN 151
RP FUNCTION, AND SUBCELLULAR LOCATION.
RX MEDLINE=99189751; PubMed=10089881;
RA Cascl T., Vinos J., Freeman M.;
RT "Sprouty, an intracellular inhibitor of Ras signaling."
RL Cell 96:655-665(1999).
CC -1- FUNCTION: INHIBITOR OF TRACHEAL BRANCHING THAT RESTRICTS BRANCH
CC BUDDING BY ANTAGONIZING THE BNL-FGF PATHWAY (BNL: BRANCHLESS, AN
CC FGF INDUCER OF BRANCHING). ACTS AS AN ANTAGONIST OF EGF-MEDIATED
CC SIGNALING IN THE EYE (WHERE IT IS IMPORTANT FOR CELL
CC DETERMINATION) MIDLINE GLIA, CHORDONAL ORGANS, WING AND OVARIAN
CC FOLLICLE CELLS.
CC -1- SUBUNIT: INTERACTS WITH DRK AND GAPI PROTEIN OF THE RAS PATHWAY.
CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE INNER SURFACE OF THE
CC PLASMA MEMBRANE.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE EMBRYONIC TRACHEAL SYSTEM,
CC DEVELOPING EYE IMAGINAL DISK, EMBRYONIC CHORDONAL ORGAN
CC PRECURSORS, MIDLINE GLIA, WING DISK AND OVARY.
CC -1- DEVELOPMENTAL STAGE: FROM STAGE 7 OF OOGENESIS FOUND IN THE
CC POSTERIOR FOLLICLE CELLS AND DURING STAGE 9 WHEN THE FOLLICLE
CC CELLS MIGRATE POSTERIORLY OVER THE OOCYTE NUCLEUS. EXPRESSION IS
CC SEEN IN THE DORSAL AND LATERAL CELLS. AND IS EXCLUDED FROM THE
CC VENTRAL CELLS. ONCE THE MIGRATION OF FOLLICLE CELLS IS COMPLETE
CC EXPRESSED IN THE DORSAL-ANTERIOR CORNER OF THE EGG CHAMBER.
CC -1- INDUCTION: BY THE BNL-FGF PATHWAY IN THE TRACHEAL SYSTEM AND BY
CC THE EGF RECEPTOR PATHWAY IN THE WING IMAGINAL DISK AND THE
CC FOLLICLE CELLS OF THE OVARY.
CC -1- DOMAIN: THE CYS-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF
CC THE PROTEIN TO THE PLASMA MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF039842; AAC04257.1; -
DR EMBL: AE003478; AAF47772.1; -
DR FlyBase: FBgn0014388; stc.
KW Developmental protein; Membrane.
FT DOMAIN 104 135 SER-RICH.
FT DOMAIN 148 162 ASN-RICH.
FT DOMAIN 207 271 GLN-RICH.
FT DOMAIN 378 501 CYS-RICH.
FT DOMAIN 492 551 GLY-RICH.
FT DOMAIN 249 262 POLY-GLN.
FT DOMAIN 503 514 POLY-GLY.
FT CONFLICT 38 N -> T (IN REF. 1).
FT CONFLICT 262 Q -> OOO (IN REF. 1).
FT CONFLICT 312 P -> L (IN REF. 1).
SQ SEQUENCE 589 AA; 62388 MW; 7450E11987B17B20 CRC64;

Query Match 23.6%; Score 391; DB 1; Length 589;
Best Local Similarity 29.3%; Pred. No. 1.9e-22;
Matches 95; Conservative 51; Mismatches 116; Indels 62; Gaps 11;

OY 4 PIPGAPITPMSYVQPLDLSRMH-SRLQHPITLPIQVKTSHVENDYIDNPSLATTT 62
DB 190 PRPESERTITNEY-----DTPLOHATRSQHPAGODNGQTTTHLLLPORNOHLHQ 243
OY 63 GPKRTGSGAPLAPTPARCDQDVTHHMFSGRPSSVSSSTSD-----QRLLDHM 115
DB 244 HQHQQ--QOOOQOOOQOOOHLQHOQNOHARLATTQATSVGSDHTDGLHSHLQNST 301
OY 116 APPYADAGSPRAVRI-----QPKVHCOPDLK----- 144
DB 302 TKPPASQKQAPRPLQMGILGILGLQNP-ITTKQTPATQKERNHALELLQPGAGCN 360
OY 145 -GPAY-----PELQKHFLCEAGCKCKECASPTPLSCVANCNECLCSQTLVNTGTC 199
DB 361 GGPLVAGDPSLNP--LVCPRCCKCEQCSPPPLPQTWVCKNTCCSAESVYDVASC 418
OY 200 MCLVQGFYHCTNED---DEGS---CADHPCSGCSRNCARMSFAGLSVVLPCLLCYL 252
DB 419 LCCAKALFEYHARDNDLDDGNGTPCVDNCSGCPYKRTQRMGLGALSTFLCLMFTW 478
OY 253 PATGCVKLAQNGYDRLRRPGRCK 276
DB 479 PMRGCKLCEKCYGRFAGRCRCQ 502

RESULT 10
SPYL_CHICK STANDARD; PRT; 76 AA.
AC Q9PTL1;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SPROUTY HOMOLOG 1 (SPRY-1) (FRAGMENT).
GN SPRT1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RX SEQUENCE FROM N.A.
RA MEDLINE=99429807; PubMed=10498682;
RA Minowada G., Jarvis L.A., Chi C.L., Neubuser A., Sun X., Hachohen N.,
RA Krasnow M.A., Martin G.R.; are induced by FGF signaling and can cause
RT chondrocyte sprouty genes are induced by FGF signaling and can cause
RL Development 126:4465-4475(1999).

CC -1- FUNCTION: MAY FUNCTION AS AN ANTAGONIST OF FIBROBLAST GROWTH
CC FACTOR (FGF) PATHWAYS AND MAY NEGATIVELY MODULATE RESPIRATORY
CC ORGANORENESIS.
CC -1- SUBCELLULAR LOCATION: FOUND IN THE CYTOPLASM IN UNSTIMULATED CELLS
CC BUT IS TRANSLOCATED TO THE MEMBRANE RUFFLES IN CELLS STIMULATED
CC WITH EGF (EPIDERMAL GROWTH FACTOR).
CC -1- TISSUE SPECIFICITY: BRAIN AND INTERLIMB REGION.
CC -1- INDUCTION: BY EGF SIGNALING.
CC -1- DOMAIN: THE CYS-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF
CC THE PROTEIN TO THE MEMBRANE RUFFLES.
CC -1- DISEASE: OVEREXPRESSION DURING SKELETAL DEVELOPMENT CAUSES A
CC SEVERE CHONDRODYSPLASIA IN WHICH CARTILAGE CELL DIFFERENTIATION IS
CC INHIBITED.
CC -1- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF177875; AAD56008.1; -
KW Developmental protein; Membrane.
FT NON_TER 1
FT DOMAIN 1 67 CYS-RICH.
SQ SEQUENCE 76 AA; 8349 MW; 6B141142FC24C2F8 CRC64;

Query Match 15.1%; Score 250; DB 1; Length 76;
Best Local Similarity 57.9%; Pred. No. 1e-12;
Matches 44; Conservative 13; Mismatches 15; Indels 4; Gaps 1;

OY 222 HPCSGSRNCARMSFAGLSVVLPCLLCYLPATGCVKLAQNGYDRLRRPGRCKHNSV 281
DB 1 NPCSGSHCCSRILCMGANSFLPCLCTPAPKGLKCGCTDVRNRRGCRCKNSNTV 60
OY 282 ICKAAS---GDARTS 293
DB 61 YCKLESPSRGGRPS 76

RESULT 11
ZAN_PIG STANDARD; PRT; 2476 AA.
AC Q28983;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ZONADHESIN PRECURSOR.
GN ZAN.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RX SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=MEISHAN; TISSUE=Testis;
RA MEDLINE=96064658; PubMed=7592795;
RA Hardy D.M., Garbers D.L.;
RT "A sperm membrane protein that binds in a species-specific manner to
RT the egg extracellular matrix is homologous to von Willebrand
RT factor".
RL J. Biol. Chem. 270:26025-26028(1995).
CC -1- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA
CC OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR
CC SIGNALING.
CC -1- SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE
CC APICAL REGION OF THE SPERM HEAD (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: IN TESTIS, PRIMARILY IN HAPLOID SPERMATIDS.
CC NOT IN LUNG, LIVER, HEART, SPLEEN, BRAIN, KIDNEY, EPIDIDYMIS.

Query Match	7.2%	Score 120:	DB 1:	Length 2476:
Best Local Similarity	19.8%	Pred. No. 0.18:	Indels 110:	Gaps 177:
Matches 63:	Conservative 33:	Mismatches 112:		
QY	2	EPPIQASAPLPNSVYVQPLDLSRMSHSLQHP-L-TTLPIDQVTSHEVDYIDNPSLA	59	
Db	552	KPTVTEKPSVTEKXPVTEPTETPTETKLVPTERTTPTKRTTPTPTLRT--TTPTR	608	
QY	60	LTTGKRTRGCAPELAPPARCDDQVYTHMISFSRPSVSSSSSTSDQRLDHMAAPP	119	
Db	609	TTTTPTERT-----TTPTRTTTTPTER-TTTPKTKTTPTEKTIPTERTI-----AP	654	
QY	120	VADQASPAVRVROPKVV-----HCQPLDLGKPAVPPELDHFLCENC	162	
Db	655	TTPOGSPPLVPQPAVAVMPSATSATVPTRTIASCPP-----NMHF-----	696	
QY	163	GKCKCK-ECASPRTPSPCMV-CNQCILC-----SAQTLVNTGTCMCLVOGIFCTNED	214	
Db	697	FKACAPVSCQSP--TPNCELFCKPCQCDVDPGLFSGSHCVNASDCDFYNDVYRLGTDW	754	
QY	215	DEGSCADHPQSC---SRNSNC-----CA-----	233	
Db	755	FSPNCTEH-CHCRPSRKMCCOTFKCGTHTVCLKNKGQCHPYGSATCSYVDDPHYLTED	813	
QY	234	--RWSFMGALSYYL--PC	247	
Db	814	GRFNFMGCKCYTLAOPC	831	
RESULT 12				
PAT3_CAEBL		STANDARD:	PRT:	809 AA.
AC	Q27874;			
DT	20-AUG-2001 (Rel. 40, Created)			
DT	20-AUG-2001 (Rel. 40, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	INTEGRIN BETA PAT-3 PRECURSOR.			
GN	PAT-3 OR ZK1058.2.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoa; Chromadorea; Rhabditida; Rhabditoidea;			
OC	Rhabditidae; Peloderinae; Caenorhabditis.			
OX	NCBI_Taxid=6239;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BRISTOL N2;			
RX	MEDLINE=9526367; PubMed=7744961;			
RA	Gettner S. N., Kenyon C., Reichardt L. F.;			
RT	"Characterization of beta pat-3 heterodimers, a family of essential			
RT	Integrin receptors in C. elegans."			
RL	J. Cell Biol. 129:1127-1141(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BRISTOL N2;			
RA	Mortimore B.;			
RL	Submitted (Aug-1994) to the EMBL/Genbank/DBJ databases.			
CC	-I- FUNCTION: POSSIBLE ROLE IN CELL-CELL INTERACTIONS. INTEGRIN ALPHA			
CC	PAT-1/BETA PAT-3 IS A RECEPTOR FOR LAMININ. INTEGRIN ALPHA PAT-			
CC	2/BETA PAT-3 RECOGNIZES THE SEQUENCE R-G-D IN ITS LIGANDS (BY			

```

CC SIMILARITY).
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA PAT-3
CC ASSOCIATES WITH EITHER ALPHA PAT-1 OR ALPHA PAT-2.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U19744; AA85704.1; -
DR EMBL: Z35604; CA84677.1; -
DR Mornep: ZK1058.2; CE01102.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002369; Integrin_B.
DR InterPro: IPR001169; Integrin_beta_C.
DR InterPro: IPR003659; PSI.
DR InterPro: IPR002035; VWA.
DR Pfam: PF00362; Integrin_B_1.
DR PRINTS: PR01186; INTEGRINB.
DR ProDom: PD001811; Integrin_B_1.
DR SMART: SM00181; EGF_1.
DR SMART: SM00187; INB_1.
DR SMART: SM00423; PSI_1.
DR SMART: SM00327; VWA_1.
DR PROSITE: PS00243; INTEGRIN_BETA_2.
DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
KW Integrin, Cell adhesion, Receptor, Glycoprotein, Transmembrane;
KW Signal, Extracellular matrix, Cytoskeleton.
FT SIGNAL 1
FT CHAIN 20 809
FT DOMAIN 20 737
FT TRANSMEM 738 758
FT DOMAIN 759 809
FT CARBOHYD 47 47
FT CARBOHYD 141 141
FT CARBOHYD 269 269
FT CARBOHYD 373 373
FT CARBOHYD 400 400
FT CARBOHYD 530 530
FT CARBOHYD 672 672
FT CARBOHYD 693 693
FT CARBOHYD 721 721
SQ SEQUENCE 809 AA; 90137 MW; 70C4AB01C8FE9189 CRC64;

Query Match 6.8%; Score 113; DB 1; Length 809;
Best Local Similarity 26.7%; Pred. No. 0.21;
Matches 27; Conservative 11; Mismatches 37; Indels 26; Gaps 5;

QY 162 CGKCKCK-----EASPTPLPS-----CWVNOECLESAQTLVNYGTCMC----- 201
DB 496 CGVCRCKGGNGYKXCECNRPMSALNENKCRKRNESALICGREGVNCNCGRCENPRANPE 555
QY 202 -LVGGIFHCNTN---EDDEGSCADH-PCSCSRNSNCARNS 236
DB 556 EQLSGFECECDNFCNPRHDKRICAEHGECNCGKCIACPGMT 596

RESULT 13
ID VWF_HUMAN STANDARD; PRT; 2813 AA.
AC P04275;
DC 20-MAR-1987 (Rel. 04, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE VON WILLEBRAND FACTOR PRECURSOR (VWF).
FMVWF OR VWF.

```

```

OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=90062044; PubMed=2584182;
RA Mancuso D.J., Tuley E.A., Westfield L.A., Morrall N.K.,
RA Shelton-Inioles B.B., Sorace J.M., Alevy Y.G., Sadler J.E.;
RT "Structure of the gene for human von Willebrand factor.";
RL J. Biol. Chem. 264:19514-19527(1989).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=87016349; PubMed=3489923;
RA Bonthron D., Orr E.C., Miltsook L.M., Ginsburg D., Handin R.I.,
RA Orlin S.H.;
RT "Nucleotide sequence of pre-pro-von Willebrand factor cDNA.";
RL Nucleic Acids Res. 14:7125-7128(1986).
RN [3]
RP SEQUENCE OF 1-120 FROM N.A., AND SEQUENCE OF 23-56.
RC TISSUE-Umbilical vein endothelial cells;
RX MEDLINE=87213253; PubMed=3495266;
RA Shelton-Inioles B.B., Broze G.J. Jr., Wilelich J.P., Sadler J.E.;
RT "Evolution of human von Willebrand factor: cDNA sequence and
RT polymorphisms, repeated domains, and relationship to von Willebrand
RT antigen II.";
RL Biochem. Biophys. Res. Commun. 144:657-665(1987).
RN [4]
RP SEQUENCE OF 1-1400 FROM N.A.
RX MEDLINE=87004550; PubMed=3019665;
RA Verweij C.L., Diegaarde P.J., Hart M., Pannekoek H.;
RT "Full-length von Willebrand factor (vWF) cDNA encodes a highly
RT repetitive protein considerably larger than the mature vWF subunit.";
RL EMBO J. 5:1839-1847(1986).
RN [5]
RP ERRATUM.
RA Verweij C.L., Diegaarde P.J., Hart M., Pannekoek H.;
RL EMBO J. 5:3074-3074(1986).
RN [6]
RP SEQUENCE OF 764-2813.
RX MEDLINE=86269895; PubMed=3524673;
RA Titani K., Kumar S., Takio K., Ericsson L.H., Wade R.D., Ashida K.,
RA Walsh K.A., Chopok M.W., Sadler J.E., Fujikawa K.;
RT "Amino acid sequence of human von Willebrand factor.";
RL Biochemistry 25:3171-3184(1986).
RN [7]
RP SEQUENCE OF 781-1424 FROM N.A.
RX MEDLINE=86269894; PubMed=3488076;
RA Shelton-Inioles B.B., Titani K., Sadler J.E.;
RT "cDNA sequences for human von Willebrand factor reveal five types of
RT repeated domains and five possible protein sequence polymorphisms.";
RL Biochemistry 25:3164-3171(1986).
RN [8]
RP SEQUENCE OF 764-873 AND 1289-2813 FROM N.A.
RX MEDLINE=86016708; PubMed=2864688;
RA Sadler J.E., Shelton-Inioles B.B., Sorace J.M., Harlan J.M.,
RA Titani K., Davie E.W.;
RT "Cloning and characterization of two cDNAs coding for human von
RT Willebrand factor.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:6394-6398(1985).
RN [9]
RP SEQUENCE OF 990-1947 FROM N.A.
RX MEDLINE=91105089; PubMed=1988024;
RA Mancuso D.J., Tuley E.A., Westfield L.A., Lester-Mancuso T.L.,
RA Le Beau M.M., Sorace J.M., Sadler J.E.;
RT "Human von Willebrand factor gene and pseudogene: structural analysis
RT and differentiation by polymerase chain reaction.";
RL Biochemistry 30:253-269(1991).
RN [10]
RP SEQUENCE OF 2731-2813 FROM N.A.
RX MEDLINE=85269603; PubMed=3875078;
RA Verweij C.L., de Vries C.J.M., Distel B., van Zonneveld A.-J.,
RA Geurts van Kessel A., van Mourik J.A., Pannekoek H.;

```



```

FT DISULFID 789 799 BY SIMILARITY.
FT DISULFID 806 817 BY SIMILARITY.
FT DISULFID 811 826 BY SIMILARITY.
FT DISULFID 828 837 BY SIMILARITY.
FT DISULFID 844 855 BY SIMILARITY.
FT DISULFID 849 890 BY SIMILARITY.
FT DISULFID 892 901 BY SIMILARITY.
FT DISULFID 908 919 BY SIMILARITY.
FT DISULFID 913 928 BY SIMILARITY.
FT DISULFID 930 939 BY SIMILARITY.
FT DISULFID 946 957 BY SIMILARITY.
FT DISULFID 952 966 BY SIMILARITY.
FT DISULFID 968 977 BY SIMILARITY.
FT DISULFID 984 995 BY SIMILARITY.
FT DISULFID 989 1009 BY SIMILARITY.
FT DISULFID 1011 1020 BY SIMILARITY.
FT DISULFID 1211 1222 BY SIMILARITY.
FT DISULFID 1216 1231 BY SIMILARITY.
FT DISULFID 1233 1242 BY SIMILARITY.
FT DISULFID 1485 1496 BY SIMILARITY.
FT DISULFID 1490 1505 BY SIMILARITY.
FT DISULFID 1507 1516 BY SIMILARITY.
FT DISULFID 1763 1774 BY SIMILARITY.
FT DISULFID 1768 1783 BY SIMILARITY.
FT DISULFID 1785 1794 BY SIMILARITY.
FT DISULFID 1801 1812 BY SIMILARITY.
FT DISULFID 1806 1821 BY SIMILARITY.
FT DISULFID 1823 1832 BY SIMILARITY.
FT DISULFID 1839 1850 BY SIMILARITY.
FT DISULFID 1844 1859 BY SIMILARITY.
FT DISULFID 1861 1870 BY SIMILARITY.
FT DISULFID 1878 1889 BY SIMILARITY.
FT DISULFID 1883 1903 BY SIMILARITY.
FT DISULFID 1905 1914 BY SIMILARITY.
FT DISULFID 1919 1930 BY SIMILARITY.
FT DISULFID 1924 1939 BY SIMILARITY.
FT DISULFID 1941 1950 BY SIMILARITY.
FT DISULFID 1957 1968 BY SIMILARITY.
FT DISULFID 1962 1977 BY SIMILARITY.
FT DISULFID 1979 1988 BY SIMILARITY.
FT DISULFID 1995 2008 BY SIMILARITY.
FT DISULFID 2002 2017 BY SIMILARITY.
FT DISULFID 2019 2028 BY SIMILARITY.
FT CARBOHYD 37 37 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 550 550 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 565 565 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 736 736 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 746 746 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 860 860 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 884 884 N-LINKED (GLCNAC. .) (POTENTIAL).

```

Query Match 6.5%; Score 108; DB 1; Length 2139;

Best Local Similarity 25.8%; Pred. No. 1.3;

Matches 63; Conservative 15; Mismatches 82; Indels 84; Gaps 15;

```

QY 68 RCGAPLAPPA-RCDD-----VTNHWISFGSPSSVSSSTSSDQ----- 109
DB 1658 KGWESWVDPPAFSTDIQAFQSLATSPQVYLGCMPEBRSQARSTLSAQQSGSFQGV 1717
QY 110 ---RLDHPMP-PPVADQAPRAVRIOPKYHOPDLKRPAYPELDKFLICEACGKC 165
DB 1718 GEARLGLLLPYFSMALVSRITVSVQKA-----QFRLNATRPBEG-----CILEFOS 1766
QY 166 KCKE---CASPRILPSCWVC-----NOECILCSAGTLVNYGTGCMCLVYG 205
DB 1767 DCKNDGCGQSPSDEYAC-TCQPGFEGDGDCTDDECLNTECL-----NNGTICINQVAA 1818

```

```

QY 206 IF-----YHCNNEDEGSCADHPSCSRNSNC-----CARNSFGALSVLP 246
DB 1819 FFCQCGPFGEGHCEQNIDE--CADQPCN-NGCNCTDLIASVYCQPE-DYWGPCQVYLK 1874
QY 247 CLIC 250
DB 1875 QMTC 1878

```

RESULT 15

BAR3_CHITE STANDARD; PRT: 1700 AA.

ID BAR3_CHITE

AC 003376;

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 01-OCT-1994 (Rel. 30, Last annotation update)

DE BALBIANT RING PROTEIN 3 PRECURSOR.

GN BR3.

OS Chironomus tentans (Midge).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;

OC Chironomidae; Chironomidae; Chironominae; Chironomus.

OX NCBI_TaxId=7153;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Salivary gland;

RX MEDLINE=90172404; PubMed=1689777;

RA Paulsson G., Lendahl U., Galli J., Ericsson C., Wieslander L.;

RT The Balbiani ring 3 gene in Chironomus tentans has a diverged repetitive structure split by many introns.*;

RL J. Mol. Biol. 211:331-349(1990).

CC -!- FUNCTION: USED BY THE LARVAE TO CONSTRUCT A SUPRAMOLECULAR STRUCTURE, THE LARVAL TUBE. BALBIANT RING PROTEIN 3 COULD PLAY A ROLE AS A TRANSPORT PROTEIN THAT BINDS TO OTHER PROTEINS INTRACELLULARLY AND IN THE GLAND LUMEN IN ORDER TO PREVENT THESE FROM FORMING WATER-INSOLUBLE FIBERS TOO EARLY.

CC -!- SUBCELLULAR LOCATION: SECRETED.

CC -!- TISSUE SPECIFICITY: SALIVARY GLAND.

CC -!- DOMAIN: HAS 82 APPROXIMATE REPEATS OF CYS-X-CYS-X-CYS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib.ch).

CC -----

DR EMBL: X52263; CA36506.1; -.

DR PIR: S08167; S08167.

DR HSSP: P18055; 2MR8.

DR InterPro: IPR000853; Multithion_nemat.

DR PRINTS: PR00876; MTNEMATODE.

KW Repeat; Signal.

FT SIGNAL 1 ?

FT CHAIN 1 ?

FT SEQUENCE 1700 AA; 186145 MW; 34202B28521B0815 CRC64;

Query Match 6.5%; Score 107.5; DB 1; Length 1700;

Best Local Similarity 27.3%; Pred. No. 1.1;

Matches 44; Conservative 17; Mismatches 49; Indels 51; Gaps 14;

```

QY 138 CQPLDKGPAVPELDKHLICEAC-----KCKKCKCASPRLPSC-----WYC 182
DB 49 CYP1---GAIFPE-SCITSMCKSGCSAKQJFNKDKAC-BCRDQPRDGGAGAGRYW-C 102
QY 183 NOECILCSAGTLVNYGTGCMCLVYGIEYHCNNEDEGSCADHPSCSRNSNCAR--WSEFG 239
DB 103 NQDCSCCKSTPMSAGCC---SGSLIWC-----EKSCA---CVCNADKCTAPQVWNNQDT 150
QY 240 ALSVPLCLCYL-----PATGCVKLAQRYGIDRLRRGCGCK 276

```

Sat May 11 12:51:27 2002

us-10-082-902-2.rsp

Page 13

Db 151 C-----CCGCPVNMQEPADGCTK--PLIMDKV---DCRCE 180

Search completed: May 8, 2002, 10:49:52
Job time: 189 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2002, 02:17:35 ; Search time 2645.75 Seconds

(without alignments)
13590.551 Million cell updates/sec

Title: US-10-082-902-1

Perfect score: 1662

Sequence: 1 aggtacgagctgagctgac.....tgtcttggaagtgtaga 1662

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 21979536 segs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents, NA, Main:*

1: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*

Result No.	Score	Query Match	Length	ID	Description
1	1662	100.0	1662	US-09-370-398-1	Sequence 1, Appl1
2	1662	100.0	1662	US-10-083-215-1	Sequence 1, Appl1
3	1662	100.0	1662	US-10-090-190-1	Sequence 1, Appl1
4	1640	98.7	2061	US-09-814-354-2	Sequence 1, Appl1
5	1324.8	79.7	1507	PCR-US01-277604-373	Sequence 373, App
6	1324.8	79.7	1507	PCR-US01-277604-373	Sequence 373, App
7	1324.8	79.7	1537	US-09-687-527-407	Sequence 407, App
8	1203.8	72.4	1208	US-09-471-275-5770	Sequence 5770, App
9	889.6	53.5	1402	US-09-360-207-6794	Sequence 6794, App
10	688.8	41.4	901	US-09-620-392-36647	Sequence 36647, App
11	675.8	40.7	680	US-09-489-036-44213	Sequence 34213, App
12	675.8	40.7	680	US-09-489-036-44213	Sequence 34213, App
13	658	39.6	658	US-09-457-877-66	Sequence 66, Appl
14	658	39.6	658	US-09-457-877-66	Sequence 66, Appl
15	583.6	35.1	635	US-09-489-036-15011	Sequence 15011, App
16	583.6	35.1	635	US-09-489-036-15011	Sequence 15011, App
17	580.6	34.9	634	US-09-489-036-14210	Sequence 14210, App
18	580.6	34.9	634	US-09-489-036-14210	Sequence 14210, App
19	569.4	34.3	634	US-09-489-036-14209	Sequence 14209, App
20	569.4	34.3	634	US-09-489-036-14209	Sequence 14209, App
21	561.6	33.8	635	US-09-943-143-34209	Sequence 34209, App
22	561.6	33.8	635	US-09-943-143-34209	Sequence 34209, App
23	476.4	28.7	546	US-09-607-200-7267	Sequence 7267, App
24	476.4	28.7	546	US-09-607-200-7267	Sequence 7267, App
25	466.2	28.1	559	US-09-649-163-10408	Sequence 10408, App
26	466.2	28.1	559	US-09-649-163-10408	Sequence 10408, App
27	443.4	26.7	449	US-09-815-264-88947	Sequence 88947, App
28	443.4	26.7	449	US-09-815-264-88947	Sequence 88947, App
29	421.6	25.4	429	US-09-311-261-2444	Sequence 2444, App
30	421.6	25.4	429	US-09-489-036-986	Sequence 986, App
31	417	25.1	423	US-09-943-143-34209	Sequence 34209, App

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES


```

Db 1501 tccattgtcttggcaccaccagagctagcctgggaggaactgcttaactgatacc 1560
Oy 1561 agggtaaccttaagaagaaccttggagctctctatactcttctctccctccca 1620
Db 1561 agggtaaccttaagaagaaccttggagctctctatactcttctctccctccca 1620
Oy 1621 ctcacacccacttctccctgatagtctctggggaaggtgtaga 1662
Db 1621 ctcacacccacttctccctgatagtctctggggaaggtgtaga 1662

RESULT 2
US-10-083-215-1
; Sequence 1, Application US/10083215
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Montgomerie, Julie R.
; TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
; FILE REFERENCE: 28110/35878
; CURRENT APPLICATION NUMBER: US/10/083, 215
; CURRENT FILING DATE: 2002-02-26
; PRIOR APPLICATION NUMBER: US/09/370,398
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (211)..(1107)
; US-10-083-215-1

Query Match 100.0%; Score 1662; DB 37; Length 1662;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 481 tccctcagcggcgcccccagctctgtgagcagcagcagcacaactcctctgaccaaag 540
Oy 541 ctcttaagcacaatggcaccaccccgctggctgacaaagcctcaccaaagcgtctgc 600
Db 541 ctcttaagcacaatggcaccaccccgctggctgacaaagcctcaccaaagcgtctgc 600
Oy 601 atccagcccaaggtgtctcaactctgacagccgctgagactcaaggcccgctgctccacc 660
Db 601 atccagcccaaggtgtctcaactctgacagccgctgagactcaaggcccgctgctccacc 660
Oy 661 gagcttgacaaagcactctctctgtctgagagcgtctgtgggaagttaaatgcaagaggt 720
Db 661 gagcttgacaaagcactctctctgtctgagagcgtctgtgggaagttaaatgcaagaggt 720
Oy 721 gcatcccccggaggttgacctctctctgctgggtctgcaaccagaggttacctgtctgacc 780
Db 721 gcatcccccggaggttgacctctctctgctgggtctgcaaccagaggttacctgtctgacc 780
Oy 781 cagactctgtcaactatgagcagctgcatgtgtttgtgtgacagggcactctctaccactgc 840
Db 781 cagactctgtcaactatgagcagctgcatgtgtttgtgtgacagggcactctctaccactgc 840
Oy 841 acgaatgagagcagatgagaggtctctctgctgacacacccctgtctctctccgctccaac 900
Db 841 acgaatgagagcagatgagaggtctctctgctgacacacccctgtctctctccgctccaac 900
Oy 901 tgcctggccggctgtgctctctcaatgggtgctctctcgtgtgtgtgtgtgtgtgtgtgtgt 960
Db 901 tgcctggccggctgtgctctctcaatgggtgctctctcgtgtgtgtgtgtgtgtgtgtgtgt 960
Oy 961 taactgtctgccaacagcgtctgtgaaagctgagccagcgtgtgtctacagacgtctctgcgcgc 1020
Db 961 taactgtctgccaacagcgtctgtgaaagctgagccagcgtgtgtctacagacgtctctgcgcgc 1020
Oy 1021 cctgtgtgctgctgtcaagcacaacagcgtctcaatgtgtgtgtgtgtgtgtgtgtgtgtgt 1080
Db 1021 cctgtgtgctgctgtcaagcacaacagcgtctcaatgtgtgtgtgtgtgtgtgtgtgtgtgt 1080
Oy 1081 aagacacagagagcccgcaaaagcctctctgacaggttgtgtgtgtgtgtgtgtgtgtgtgt 1140
Db 1081 aagacacagagagcccgcaaaagcctctctgacaggttgtgtgtgtgtgtgtgtgtgtgtgt 1140
Oy 1141 tggaaacctgtgtctctctctgacatctaagaagactgacgaagaagtgcaagggtttagcc 1200
Db 1141 tggaaacctgtgtctctctctgacatctaagaagactgacgaagaagtgcaagggtttagcc 1200
Oy 1201 tccctgaggtctgacctgtgctagctgtgcccactcctaccacccagctctggaataacagag 1260
Db 1201 tccctgaggtctgacctgtgctagctgtgcccactcctaccacccagctctggaataacagag 1260
Oy 1261 accacacacacgtaccctgtgtctctcccaagaatgtaagaagaagcacttggggcttttt 1320
Db 1261 accacacacacgtaccctgtgtctctcccaagaatgtaagaagaagcacttggggcttttt 1320
Oy 1321 cagggtccctgaacttgtgtcaaacagacatgcaagggcgaaggtgtgtgtgtgtgtgtgtgt 1380
Db 1321 cagggtccctgaacttgtgtcaaacagacatgcaagggcgaaggtgtgtgtgtgtgtgtgtgt 1380
Oy 1381 aatttctctttttagaagaagacagaacacagatgtgtgacacatataccggaactgtcagctg 1440
Db 1381 aatttctctttttagaagaagacagaacacagatgtgtgacacatataccggaactgtcagctg 1440
Oy 1441 ctgtgaatgctctcccaagccctctctctctccctccctccctccctccctccctccctccct 1500
Db 1441 ctgtgaatgctctcccaagccctctctctctccctccctccctccctccctccctccctccct 1500
Oy 1501 tccattgtcttggcactacacagagctagctgctggggaagtgtctaactgagtaacc 1560
Db 1501 tccattgtcttggcactacacagagctagctgctggggaagtgtctaactgagtaacc 1560
Oy 1561 agggtaaccttaagaagaaccttggagctctctatactcttctctcccatctca 1620

```

QY	1	aggtacagcagctctgagctctgacacgacgagctcgtgacgctctgagagcagggagacctctgcg	60
Db	1	aggtacagcagctctgagctctgacacgacgagctcgtgacgctctgagagcagggagacctctgcg	60
QY	61	gttcctccggaacggagctctgagagtaaaagcggcggctctaacctctgcccggctctcagagattc	120
Db	61	gttcctccggaacggagctctgagagtaaaagcggcggctctaacctctgcccggctctcagagattc	120
QY	121	acacagacgctgaggcagatctgtctgtgacctgagctctctcaaaagcgcctcaagaagcctgtc	180
Db	121	acacagacgctgaggcagatctgtctgtgacctgagctctctcaaaagcgcctcaagaagcctgtc	180
QY	181	tctccgcgtacagctccaaagacctccaaagcccaatggagcccccagatcccaacagagagccccc	240
Db	181	tctccgcgtacagctccaaagacctccaaagcccaatggagcccccagatcccaacagagagccccc	240
QY	241	ttagctcccaacatcagatctgctctgacagccctctctctgacagccggatctcccaacagccgg	300
Db	241	ttagctcccaacatcagatctgctctgacagccctctctctgacagccggatctcccaacagccgg	300
QY	301	ctccagacacccatcacaacatctctcaacatctgaccaagctgaagacacagcatctgtgagagat	360
Db	301	ctccagacacccatcacaacatctctcaacatctgaccaagctgaagacacagcatctgtgagagat	360
QY	361	gactacataagcaaacccctcagctctgacctgacacacggcccaaaagcgagaccccggtggcg	420
Db	361	gactacataagcaaacccctcagctctgacctgacacacggcccaaaagcgagaccccggtggcg	420
QY	421	ggcccaagctcgtggcccgagcccgcccgccgtctgtaaccagagatgtctcaaccacatcttgatc	480
Db	421	ggcccaagctcgtggcccgagcccgcccgccgtctgtaaccagagatgtctcaaccacatcttgatc	480
QY	481	tctcttcagcggcgcccccagctctctgtgagacagcagcagcaatactctctgacacaacgg	540
Db	481	tctcttcagcggcgcccccagctctctgtgagacagcagcagcaatactctctgacacaacgg	540
QY	541	ctctttagaccatctgacacacacccgttgctctgtaaccagacctcaacaaaggtctgtgcg	600
Db	541	ctctttagaccatctgacacacacccgttgctctgtaaccagacctcaacaaaggtctgtgcg	600

Dh	541	ctctttagacacacatgagccacacacaccccgctgctctgacacagctccatccacaaagggctctgtgcgc	600
Qy	601	atccagagccccaaggtgtgtctccactgtccagccgctgtgagactccataagggcccggtgtccacacc	660
Dh	601	atccagagccccaaggtgtgtctccactgtccagccgctgtgagactccataagggcccggtgtccacacc	660
Qy	661	gaagcttgagacaaagacactctctgtctgtctgaggggcctgtgtggaaagtgtaaaatgcaaggaagtgt	720
Dh	661	gaagcttgagacaaagacactctctgtctgtctgaggggcctgtgtggaaagtgtaaaatgcaaggaagtgt	720
Qy	721	gcatctcccccagaaagtgtgctctctccctctccgtgtctgaaacaaagaaatgtgactgtgtccaaagc	780
Dh	721	gcatctcccccagaaagtgtgctctctccctctccgtgtctgaaacaaagaaatgtgactgtgtccaaagc	780
Qy	781	cagaacctctgtgtcacaactctatgagacgtgacgtgtttgtgtgcaggtgacatctctccataccatgtgc	840
Dh	781	cagaacctctgtgtcacaactctatgagacgtgacgtgtttgtgtgcaggtgacatctctccataccatgtgc	840
Qy	841	aaggaatagagagacgtatgtagggctctccctgtcgtctgaaacacccctgtctctctgtcccgctccacac	900
Dh	841	aaggaatagagagacgtatgtagggctctccctgtcgtctgaaacacccctgtctctctgtcccgctccacac	900
Qy	901	tgctgtgccccccgtctgtatccctctcatctgtgtgtgtctctccgtgtgtctccctctgctcgtcgtgc	960
Dh	901	tgctgtgccccccgtctgtatccctctcatctgtgtgtgtctctccgtgtgtctccctctgctcgtcgtgc	960
Qy	961	taacctgtccctgacacacggctctgtctgtgaaagctgtgccccagcgtgtgtctacagacacgtctgcgcgc	1020
Dh	961	taacctgtccctgacacacggctctgtctgtgaaagctgtgccccagcgtgtgtctacagacacgtctgcgcgc	1020
Qy	1021	ccctgtgttgcccgctgtgacaaagacacagaaacagcgtatcatctgaaagacagccaggtggggatgtgcc	1080
Dh	1021	ccctgtgttgcccgctgtgacaaagacacagaaacagcgtatcatctgaaagacagccaggtggggatgtgcc	1080
Qy	1081	aagagccacagccaggtccccgcgaacaaagccctctctctgacagattgtgtctgaaagcccccaagtgtctctgc	1140
Dh	1081	aagagccacagccaggtccccgcgaacaaagccctctctctgacagattgtgtctgaaagcccccaagtgtctctgc	1140
Qy	1141	tgtgaaacactgtgtctctctctctgtacacatctataagaaagacatgtgcagaaagtctgaaagcttttagcc	1200
Dh	1141	tgtgaaacactgtgtctctctctctgtacacatctataagaaagacatgtgcagaaagtctgaaagcttttagcc	1200
Qy	1201	ttccctgaaagctgtgacctgtgtctgtctgtccacactccctacaccccccagctcttggaataataacagag	1260
Dh	1201	ttccctgaaagctgtgacctgtgtctgtctgtccacactccctacaccccccagctcttggaataataacagag	1260
Qy	1261	accacacacacaggtacccctgtgtatctcccccacagaaatgtatgaaagaaacacattctgtgggctttttt	1320
Dh	1261	accacacacacaggtacccctgtgtatctcccccacagaaatgtatgaaagaaacacattctgtgggctttttt	1320
Qy	1321	cagaggtctctgtaaaactctgtgtctctctctctctctctctccacacatgtgcaggggcgtgtgtgtgtgggga	1380
Dh	1321	cagaggtctctgtaaaactctgtgtctctctctctctctctctccacacatgtgcaggggcgtgtgtgtgtgggga	1380
Qy	1381	aattttctctttctcagaagaagacagaaacagaaatgtgtgaaacatatccoggaaaaactgtccagctgt	1440
Dh	1381	aattttctctttctcagaagaagacagaaacagaaatgtgtgaaacatatccoggaaaaactgtccagctgt	1440
Qy	1441	cttgaaatgacctctccacacacccctctctctctccctccctccctccctccctccctccctctctctct	1500
Dh	1441	cttgaaatgacctctccacacacccctctctctctccctccctccctccctccctccctccctccctctctct	1500
Qy	1501	tccaatgtctctgtgacactccaaagagactgtacgtccctgtgagagaaatctgtctaaactgtatgtaacc	1560
Dh	1501	tccaatgtctctgtgacactccaaagagactgtacgtccctgtgagagaaatctgtctaaactgtatgtaacc	1560
Qy	1561	agggtacacctttaaagaaagacctgtgagactctctctataacactctctctccctccacatctca	1620
Dh	1561	agggtacacctttaaagaaagacctgtgagactctctctataacactctctctccctccacatctca	1620
Qy	1621	ctccacacccacattgtctccctcgtatgactctgtggaggaagtgtgaga	1662
Dh	1621	ctccacacccacattgtctccctcgtatgactctgtggaggaagtgtgaga	1662

Qy	1623 ccaacccacctgtccctcatgctcttgggaagatcaga	1662
Db	1656 caaccacctgtccctgatgtcttggaagatcaga	1695

RESULT 5
PCT-11501-3

```

: Sequence: 373, Application: PC/TUS0127760
:
: GENERAL INFORMATION:
:
:   APPLICANT: Hyseq, Inc
:
:   TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
:   FILE REFERENCE: 21272-115/ 785
:   CURRENT APPLICATION NUMBER: PCT/US01/27760
:   CURRENT FILING DATE: 2001-10-11
:   PRIOR APPLICATION NUMBER: 09/667, 527
:   PRIOR FILING DATE: 2000-10-12
:   NUMBER OF SEQ. ID NOS: 892
:
:   SOFTWARE: Custom
:
:   SEQ ID NO 373
:
:   LENGTH: 1507
:
:   TYPE: DNA
:
:   ORGANISM: Homo sapiens
:
:   FEATURE:
:
:   NAME/KEY: CDS
:   LOCATION: (154)..(954)
:   CDT-US01-27760-373

```

Query Match	79.78;	Score 1324.8;	DB 1;	Length 1507;
Best Local Similarity	80.19;	Prod. No. 348-205.		

[illegible]

OY	919	ttcaabvgvgtctctcccccvgvgtgacvgtccctctgctctctgctctcaaccttgctccacccgvc	978
Db	763	ttcaatvgvgtctctctccctccgctvgvgtgctccctctgctctctctctctcaacctgctccacccgvc	822
OY	979	tgvcgtgaaagctvgvgtcccaagcgtgvtgacgtcaacgtctctgctccgctcccttgcttgccgvtgacag	1038
Db	823	tgvcgtgaaagctvgvgtcccaagcgtgvtgacgtcaacgtctctgctccgctcccttgcttgccgvtgacag	882
OY	1039	cacacgaaacagcgtctcatctctgcaaaagcagccagccagcgtggtatgccaagaaccagagcccgac	1098
Db	883	cacacgaaacagcgtctcatctctgcaaaagcagccagccagcgtggtatgccaagaaccagagcccgac	942
OY	1099	aagcccttctcgacaggtttgtgtccgaagccccaatgtgtctgctcccttgtaaaacctgtgtctctt	1158
Db	943	aagcccttctcgacaggtttgtgtccgaagccccaatgtgtctgctcccttgtaaaacctgtgtctctt	1002
OY	1159	ctgacatctcaagaaagactctgcaagacaagctccaaagttcttaagctctctctgtagagctcgaccttgc	1218
Db	1003	ctgacatctcaagaaagactctgcaagacaagctctcaaggttttagctctctctgtagagctcgaccttgc	1062
OY	1219	taagctctgcccactctcccttaacccccagctctgtaaaatataagaaaccaacacagctaacct	1278
Db	1063	taagctctgcccactctcccttaacccccagctctgtaaaatataagaaaccaacacagctaacct	1122
OY	1279	gtatctccccaagaatgatvgaaagaaagccttttgvggctcttttctcaaggtctccctgaaaaccttg	1338
Db	1123	gtatctccccaagaatgatvgaaagaaagccttttgvggctcttttctcaaggtctccctgaaaaccttg	1182
OY	1339	tgctcaaacagaaacaaatctgcaagvgcagagctgtgtgtcttgvgvggaaatctctctcttctcaagaa	1398
Db	1183	tgctcaaacagaaacaaatctgcaagvgcagagctgtgtgtcttgvgvggaaatctctctcttctcaagaa	1242
OY	1399	gtaacgaacacagatgtgtgacacatactccvggaaactgtgacgtgtgaaatgtccctccacgc	1458
Db	1243	gtaacgaacacagatgtgtgacacatactccvggaaactgtgacgtgtgaaatgtccctccacgc	1302
OY	1459	ccctctctctctccctccctctccctctccgccccccctccctccctctcttctcaatgtcttggaact	1518
Db	1303	ccctctctctctccctccctccctctccgccccccctccctccctctcttctcaatgtcttggaact	1352
OY	1519	cacacgagctctagctctgctctggaagaaatctgctcaactgagctaacccaggtatccctttaaagaag	1578
Db	1363	cacacgagctctagctctgctctggaagaaatctgctcaactgaggtaacccaggtatccctttaaagaag	1422
OY	1579	aacctgtgagatctctataactctctctctctccctcccaatccatccacccacacttgctcc	1638
Db	1423	aacctgtgagatctctataactctctctctctccctcccaatccatccacccacacttgctcc	1482
OY	1639	ctgagtctctgvggaaagtgtaga	1662
Db	1483	ctgagtctctgvggaaagtgtaga	1506

RESULT 6

```

1 Sequence: 373; Application: PC/TUS0127760A
2
3 GENERAL INFORMATION:
4
5 APPLICANT: Hysec, Inc
6
7 TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
8
9 FILE REFERENCE: 21272-115 / 785
10
11 CURRENT APPLICATION NUMBER: PCT/TUS01/27760A
12
13 PRIOR FILING DATE: 2001-10-11
14
15 PRIOR APPLICATION NUMBER: 09/687,527
16
17 PRIOR FILING DATE: 2000-10-12
18
19 NUMBER OF SEQ ID NOS: 892
20
21 SOFTWARE: Custom
22
23 SEQ ID NO 373
24
25 LENGTH: 1507
26
27 TYPE: DNA
28
29 ORGANISM: Homo sapiens
30
31 FEATURE:
32
33 NAME/KEY: CDS
34
35 LOCATION: (154)..(954)

```


DB 1088 CAGAGGCTCAGCAAGGCGTGTGGCATCCAGCCCAAGGTGTCTACATGCGACGCGCTGGA 1029
OY 636 cctcaaggccgagcggtcccaaccgagctgcaagcaacttctgctgagggcctg 695
DB 1028 CCTCAAGGCGCGGCGGTCCACCGAGCTGAGCAAGCACTTCTGCTGCGAGGCGCTG 969
OY 696 tgggaagtgaatgcaaggaagtgcatacccccgaagctgctctctgctgggtctg 755
DB 968 TGGGAAGTGAATGCAAGAGAGTGTGATCCCGGACGCTTCCTCCGCTGGGTCTG 909
OY 756 caaccaggaatgctgctgctcagcccaagactctgctcaactgagcaagtgcatt 815
DB 908 CAACGAGGATGCTGCTGTGCTCAGCCAGACTGTGCTCAATATGCGACGTGTCTT 849
OY 816 ggtgcaaggagcatcttcaacactgcaagaa tgaagcagatgaaggtcctgctgacca 875
DB 848 GGTGAGGCGATCTGTACACACTGACAGATGAGAGAGGCGCTCGCGCTGACCA 789
OY 876 cccctgtctctgtcccgctcccaactgctgagccgctgcttctataggtctctc 935
DB 788 CCCCTGCTCTGTGCTCCGCTCCACTGCTGCGCCGCTGTGCTTCATGGGTCTCTTC 729
OY 936 cgtggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 995
DB 728 CGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 669
OY 996 gctggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1055
DB 668 GCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 609
OY 1056 ctgcaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1115
DB 608 CTGCAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 549
OY 1116 tctgtcgaagcccaagctgctgctgctgctgctgctgctgctgctgctgctgctg 1175
DB 548 TGTGTGCAAGCCCCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 469
OY 1176 tgcagcaagctgcaagagcttctgctgctgctgctgctgctgctgctgctgctg 1235
DB 488 TGCAGCAAGCTGAGAGGTGTTAGGCTCTGAGGCTGACCTTGTGAGTCCGCTGCT 429
OY 1236 accccagctctggaataatacagagacacacacacacacacacacacacacacac 1295
DB 428 ACCCCAGCTTCGAAATACAGAGACACACACACACACACACACACACACACACAC 369
OY 1296 gaagaagcacttgggcttcttctgaggtcctgaaacttctgtaaacagaaatgc 1355
DB 368 GAGGAAGCAGCTTGGGGCTTTTTCAGGGTCTGAAACTTGTGTAAACAGCAATGTC 309
OY 1356 aggggagaggtggtgcttggggggaattcttcttcttcttcttcttcttcttctt 1415
DB 308 AGGGGAGGAGGTGTGTTGGGGGGAATTTTCTTTTTCAGAGACAGACAGATGTC 249
OY 1416 gaacataatcagaaactgagctgctgaa tgcctccagccctctctctctcc 1475
DB 248 GACACATATCCGGAACCTGACAGCTGTGATGCTTCCAGGCCCTCTTCCCTCC 189
OY 1476 tccctcgcccccctctctcttcttcttcttcttcttcttcttcttcttcttctt 1535
DB 188 TCCCTCCGCCCCCT 129
OY 1536 tggaggaatgctgaactgagctacagagctacacttaagaagacacttgaatcttcta 1595
DB 128 TGGGAGGAATTTGTAATGAGTACAGAGGTACCTTTAAAGAGACCTTGGAGCTTCTTA 69
OY 1596 taacttctctctctcccatctcaatccacccacttgccttga tgccttggagag 1655
DB 68 TACCTTCT 9
OY 1656 gtgtaga 1662
|||||||

DB 8 GTGTAGA 2
RESULT 9
US-60-360-207-6794
; Sequence 6794, Application US/60360207
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
; FILE REFERENCE: CLO01321
; CURRENT APPLICATION NUMBER: US/60/360,207
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 47235
; SEQ ID NO 6794
; LENGTH: 1402
; TYPE: DNA
; ORGANISM: HUMAN
US-60-360-207-6794
Query Match 53.5%; Score 889.6; DB 75; Length 1402;
Best Local Similarity 86.4%; Pred. No. 3,9e-188;
Matches 995; Conservative 0; Mismatches 154; Indels 3; Gaps 1;
OY 1 aggttagcagcgtgagctgagcagcagcagcagcagcagcagcagcagcagcagcagc 60
DB 88 agcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 147
OY 61 gtcctccacacgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 120
DB 148 gctcctccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 207
OY 121 acacagacgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 180
DB 208 acacagacgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 267
OY 181 tctcctgtaacgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 237
DB 268 ccttggctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 327
OY 238 ccttggctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 297
DB 328 ccttggctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 387
OY 298 cggctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 357
DB 388 cggctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 447
OY 358 aatgactacatagac 417
DB 448 aatgactacatagac 507
OY 418 ggggcccagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 477
DB 508 ggggcccagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 567
OY 478 atctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 537
DB 568 atctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 627
OY 538 cggctccttagacacatgagcagcagcagcagcagcagcagcagcagcagcagcagcag 597
DB 628 aggcctcagacatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 687
OY 598 cggctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 657
DB 688 cggctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 747
OY 658 cccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 717
DB 748 ccagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 807
OY 718 tctgcatcccccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 777

	Query Match	41.4%;	Score 688.8;	DB 24;	Length 901;
	Best Local Similarity	86.7%;	Pred. N.2.5e-143;		
	Matches 782;	Conservative 0;	Mismatches 117;	Indels 3;	Gaps 2;
QY	211	atggaagccccgcatccacagagcgcc--cccttgaactcccaatcagtcagtcagcgcg	268		
Db	901	atggagcccccggtgtcccaaaacacttcccttggaattccaaattcaattggggcaac	842		
QY	269	ccctctcttgaagcccgagatgtcccaacagcggtctccaaacccactatccatctccaca	328		
Db	841	ccctcttcttgaaacaaaacccccccacacgcggctctcacagaccactctacac--ttttaacca	783		
QY	329	ttgacccaggtgaagaccagcagcatgtggaagaa tgaactatagacacacctaagccttgcgc	388		
Db	782	ttggccacagatgaagaccacgcagccacgtggagaa tgaactatagacaaaccttgcctggccc	723		
QY	389	tgaacacccgggcccgaagagagagccggggcgagcccccaagcttgcgccccgagccgcgc	448		
Db	722	ctggccacagcgttcccaagagggcccccggggcgggccacccagagctgctcttcaacctgtgcc	663		
QY	449	gcctgtgaccagagatgtcacccacacat tgaatcctcttlaagcggcgcccccaactctgtga	508		

	Query Match	40.7%	Score 675.8	DB 18	Length 660
Best Local Similarity	99.7%	Pred. No. 1.9e+40			
Matches 677	Conservative	0	Mismatches 2	Indels 0	Gaps 0
QY	984	gaagctggccacagcttgctcagacagctctgcgcgccttgcttcgcgtcgaagacac	1043		

Db 1 gaagctgcccagcgtgtgtaacgacgtctgcgcgcctgtgtgcgcgtgcaagcaac 60
QY 1044 gaacagcgtcatctgcaaaagcagccagcggggaatgcaaaagcagcagccgcaaac 1103
Db 61 gaacagcgtcatctgcaaaagcagccagcggggaatgcaaaagcagcagccgcaaac 120
QY 1104 ttcttgacagttgtgtcgaaagccagctgtctgtcgtgaaacctgttctcttgac 1163
Db 121 ttcttgacagttgtgtcgaaagccagctgtctgtcgtgaaacctgttctcttgac 180
QY 1164 atctaaagagctgcaagcaggtttagcctcttgagcgtgagcctgtgac 1223
Db 181 atctaaagagctgcaagcaggtttagcctcttgagcgtgagcctgtgac 240
QY 1224 tgcacatccctaccccccagcttcggaataaagagacacacacacgtaacctgtat 1283
Db 241 tgcacatccctaccccccagcttcggaataaagagacacacacacgtaacctgtat 300
QY 1284 ccccaagatgataaagaagcacttgggctttttcaggttcctgaaacctgtgtca 1343
Db 301 ccccaagatgataaagaagcacttgggctttttcaggttcctgaaacctgtgtca 360
QY 1344 aacagcaatgacgggagcaggtgtgtgttggggggaattttctttcagaagaacag 1403
Db 361 aacagcaatgacgggagcaggtgtgtgttggggggaattttctttcagaagaacag 420
QY 1404 aacagcagtgatgacacataatccggaactgacgctgtgaaatgcttccagccctc 1463
Db 421 aacagcagtgatgacacataatccggaactgacgctgtgaaatgcttccagccctc 480
QY 1464 ctctccctccctccctccgccccctctctcttccatgtcttggcactcaag 1523
Db 481 ctctccctccctccctccgccccctctctcttccatgtcttggcactcaag 540
QY 1524 gaagtagctgctggaggaattgtaactgagatcagaggtactcttaagaagacct 1583
Db 541 gaagtagctgctggaggaattgtaactgagatcagaggtactcttaagaagacct 600
QY 1584 tggagctctatactctctctctcccatctcaatccacacacacacttgcctgat 1643
Db 601 tggagctctatactctctctctcccatctcaatccacacacacacttgcctgat 660
QY 1644 gtcttggggaaggtgtaga 1662
Db 661 gtcttggggaaggtgtaga 679

RESULT 12

US-09-943-143-34213
; Sequence 34213, Application US/09943143
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 783
; CURRENT APPLICATION NUMBER: US/09/943,143
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/489,036
; PRIOR FILING DATE: 2000-01-19
; NUMBER OF SEQ ID NOS: 35324
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 34213
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-943-143-34213

Query Match 40.7%; Score 675.8; DB 35; Length 680;
Best Local Similarity 99.7%; Pred. No. 1.9e-140;
Matches 677; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 984 gaagctggccagcgtgtgtaacgacgtctgcgcgcctgtgtgtgcgcgtgcaagcaac 1043

Db 1 gaagctggccagcgtgtgtaacgacgtctgcgcgcctgtgtgtgcgcgtgcaagcaac 60
QY 1044 gaacagcgtcatctgcaaaagcagccagcggggaatgcaaaagcagcagccgcaaac 1103
Db 61 gaacagcgtcatctgcaaaagcagccagcggggaatgcaaaagcagcagccgcaaac 120
QY 1104 ttcttgacagttgtgtcgaaagccagctgtctgtcgtgaaacctgttctcttgac 1163
Db 121 ttcttgacagttgtgtcgaaagccagctgtctgtcgtgaaacctgttctcttgac 180
QY 1164 atctaaagagctgcaagcaggtttagcctcttgagcgtgagcctgtgac 1223
Db 181 atctaaagagctgcaagcaggtttagcctcttgagcgtgagcctgtgac 240
QY 1224 tgcacatccctaccccccagcttcggaataaagagacacacacacgtaacctgtat 1283
Db 241 tgcacatccctaccccccagcttcggaataaagagacacacacacgtaacctgtat 300
QY 1284 ccccaagatgataaagaagcacttgggctttttcaggttcctgaaacctgtgtca 1343
Db 301 ccccaagatgataaagaagcacttgggctttttcaggttcctgaaacctgtgtca 360
QY 1344 aacagcaatgacgggagcaggtgtgtgttggggggaattttctttcagaagaacag 1403
Db 361 aacagcaatgacgggagcaggtgtgtgttggggggaattttctttcagaagaacag 420
QY 1404 aacagcagtgatgacacataatccggaactgacgctgtgaaatgcttccagccctc 1463
Db 421 aacagcagtgatgacacataatccggaactgacgctgtgaaatgcttccagccctc 480
QY 1464 ctctccctccctccctccgccccctctctcttccatgtcttggcactcaag 1523
Db 481 ctctccctccctccctccgccccctctctcttccatgtcttggcactcaag 540
QY 1524 gaagtagctgctggaggaattgtaactgagatcagaggtactcttaagaagacct 1583
Db 541 gaagtagctgctggaggaattgtaactgagatcagaggtactcttaagaagacct 600
QY 1584 tggagctctatactctctctctcccatctcaatccacacacacacttgcctgat 1643
Db 601 tggagctctatactctctctctcccatctcaatccacacacacacttgcctgat 660
QY 1644 gtcttggggaaggtgtaga 1662
Db 661 gtcttggggaaggtgtaga 679

RESULT 13

US-09-457-877-66/c
; Sequence 66, Application US/09457877
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Contigs Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 20411-779
; CURRENT APPLICATION NUMBER: US/09/457,877
; CURRENT FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 66
; LENGTH: 658
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (51)...(459)
; OTHER INFORMATION: similar to V32399/V32399 in the geneseq database, Run with
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (51)...(459)
; OTHER INFORMATION: similar to gb|AF039843|AF039843 in the nt_19990223 databas

```

OTHER INFORMATION: e, Run with BLASTN 2.0a19MP-WashU, default parameters
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3)...(365)
OTHER INFORMATION: similar to gb|AA499432|AA499432 in the dbEST weekly FASTA.
OTHER INFORMATION: 021899 database, Run with BLASTN 2.0a19MP-WashU, default par
OTHER INFORMATION: ameters
FEATURE:
NAME/KEY: source
LOCATION: (235)...(659)
OTHER INFORMATION: Variant: 638088, Cluster: 582068, Library: adt.brain BRAIN G
OTHER INFORMATION: Variant: 638088, Cluster: 582068, Library: adt.brain BRAIN G
FEATURE:
NAME/KEY: source
LOCATION: (1)...(452)
OTHER INFORMATION: Variant: 779832, Cluster: 717962, Library: thymus THYMUS GF/
OTHER INFORMATION: Variant: 779832, Cluster: 717962, Library: thymus THYMUS GF/
US-09-457-87-66

```

```

Query Match          39.6%; Score 658; DB 18; Length 658;
Best Local Similarity 100.0%; Pred. No. 1.8e-136;
Matches 658; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 462 tgtacacacacatttgatctctctcctcagcgggcgcccaagctctgtgagcagcagcag 521
DB 658 TGTACACACACATTGATCTCTCTCTCAGCGGGCGCCCAAGCTCTGTGAGCAGCAGCAG 599
QY 522 cacaacctctgacacagcgctctctctcagcagcagcagcagcagcagcagcagcag 581
DB 598 CACATCTCTGTGACACAGCGCTCTTATGACACATGACACACACACCGCTGCTGACACAGC 539
QY 582 ctcaacaaagggctgtgctgacatccagccaaaggtgtgtcactgtcagcagcagcagc 641
DB 538 CTCACCAAGGCGCTGTGCGCATCCAGCCCAAGGTGTGTGCTGCTGCGCAGCCCTGACCTCAA 479
QY 642 gggcgcgggcgctgacacacagcgctctcagcagcagcagcagcagcagcagcagcag 701
DB 478 GGGCGCGGGCGCTGACACACCGCTGCGCAAGCACTTCTGTGCGCAGCGCTGTGCGGAA 419
QY 702 gtgtaaatgcaagagtgatgcatcccccgcgagctgtcctcctcgtggtgtgtgtgtgt 761
DB 418 GTGTAATGCAAGAGTGTGATGCCCGCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCAACCA 359
QY 762 ggaagtcgtgtgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 821
DB 358 GGAAGTGCCTGTGCTGCAACCCAGACTGTGCTGCACTATGCAAGCATGTGTGTGTGTGCA 299
QY 822 gggcatctctacacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 881
DB 298 GGGCATCTCTTACCACTGCAAGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 239
QY 882 ctctgtctcgcgtctcacaactgtcgcgcgcgcgtgtgtgtgtgtgtgtgtgtgtgtgt 941
DB 238 CTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 179
QY 942 gctgcgcctgtcgtctgtcactgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1001
DB 178 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 119
QY 1002 ctacgacgcgtctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1061
DB 118 CTACGACGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 59
QY 1062 agagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1119
DB 58 AGCAGCAGCAGCGGGGATGCGCAAGACAGCAGCGCCGCAAGCCTTTCTGTGACAGTTTGTG 1

```

```

RESULT 14
US-09-952-981-66/c
; Sequence 66, Application US/09952981
; GENERAL INFORMATION:

```

```

APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: Novel Configs Obtained
FILE REFERENCE: 20411-779
CURRENT APPLICATION NUMBER: US/09/952,981
PRIOR FILING DATE: 2001-09-14
PRIOR FILING DATE: 1999-12-08
SOFTWARE: Hy-patent.pl Version 3.1
SEQ ID NO 66
LENGTH: 658
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51)...(459)
OTHER INFORMATION: similar to V32399|V32399 in the geneseq database, Run
OTHER INFORMATION: BLASTN 2.0a19MP-WashU, default parameters
NAME/KEY: misc_feature
LOCATION: (51)...(459)
OTHER INFORMATION: similar to gb|AF039843|AF039843 in the nt_19990223
OTHER INFORMATION: dbEST weekly FASTA.
OTHER INFORMATION: 021899 database, Run with BLASTN 2.0a19MP-WashU, default par
OTHER INFORMATION: ameters
NAME/KEY: source
LOCATION: (235)...(659)
OTHER INFORMATION: Variant: 638088, Cluster: 582068, Library: adt.brain
OTHER INFORMATION: Variant: 638088, Cluster: 582068, Library: adt.brain
OTHER INFORMATION: BRAIN G
OTHER INFORMATION: Ienn Fu
NAME/KEY: source
LOCATION: (1)...(452)
OTHER INFORMATION: Variant: 779832, Cluster: 717962, Library: thymus
OTHER INFORMATION: THYMUS GF/
OTHER INFORMATION: MW
US-09-952-981-66

```

```

Query Match          39.6%; Score 658; DB 35; Length 658;
Best Local Similarity 100.0%; Pred. No. 1.8e-136;
Matches 658; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 462 tgtacacacacatttgatctctcctcagcgggcgcccaagctctgtgagcagcagcag 521
DB 658 TGTACACACACATTGATCTCTCTCTCAGCGGGCGCCCAAGCTCTGTGAGCAGCAGCAG 599
QY 522 cacaacctctgacacagcgctctctcagcagcagcagcagcagcagcagcagcagcagcag 581
DB 598 CACATCTCTGTGACACAGCGCTCTTATGACACATGACACACACCGCTGCTGACACAGC 539
QY 582 ctcaacaaagggctgtgctgacatccagccaaaggtgtgtcactgtcagcagcagcagcag 641
DB 538 CTCACCAAGGCGCTGTGCGCATCCAGCCCAAGGTGTGTGCTGCTGCGCAGCCCTGACCTCAA 479
QY 642 gggcgcgggcgctgacacacagcgctctcagcagcagcagcagcagcagcagcagcag 701
DB 478 GGGCGCGGGCGCTGACACACCGCTGCGCAAGCACTTCTGTGCGCAGCGCTGTGCGGAA 419
QY 702 gtgtaaatgcaagagtgatgcatcccccgcgagctgtcctcctcgtggtgtgtgtgtgt 761
DB 418 GTGTAATGCAAGAGTGTGATGCCCGCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCAACCA 359
QY 762 ggaagtcgtgtgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 821
DB 358 GGAAGTGCCTGTGCTGCAACCCAGACTGTGCTGCACTATGCAAGCATGTGTGTGTGTGCA 299
QY 822 gggcatctctacacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 881

```

[illegible]

```

RESULT 15
US-09-489-036-15011/C
: Sequence 15011, Application US/09489036
: GENERAL INFORMATION:
: APPLICANT: Hyseq, Inc.
: TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
: TITLE OF INVENTION: From Various Libraries
: FILE REFERENCE: 783
: CURRENT APPLICATION NUMBER: US/09/489,036
: CURRENT FILING DATE: 2000-01-19
: NUMBER OF SEQ ID NOS: 35324
: SOFTWARE: Hy-Patent.pl Version 3.1
: SEQ ID NO 15011
: LENGTH: 635
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-489-036-15011

```

Query Match	35.1%;	Score 583.6;	DB 18;	Length 635;
Best Local Similarity	98.7%;	Pred. NO. 7.4e-120;		
Matches 620;	Conservative 0;	Mismatches 4;	Indels 4;	Gaps 3

QY	53	ccctggccgattcccccga - ccggagctctgaaatagaacggcgagctaaactgtcccccgt	111
Db	626	CTTtgcCGGTTCTCTCCAGCCGGGCTTGGGAGTACAGCGGGGCTTAACTGGCCCGGCT	567
QY	112	tcaggaattacacagacgttgggcgatgcttg - accctgcagctctctaaagccccc	170
Db	566	TCAGAGATTACAGAGAGCTGGGGGAGTGGCTTTGTAACCCtGCAGtCTCTTAAAGGCCCT	507
QY	171	agaagcctgtttctcgtacagtcacagaccctccagcccaatggaagccccgaatccca	230
Db	506	AGAAGCCTGTtTCTCCGTACAGTCCAGAGACTCCAGCCCATGGAGGCCCGATCCACA	447
QY	231	gagagcccccctgtgctccaaactaagtcagatggtctccagcccccctctgtacagccgagatgc	290
Db	446	GAGGCCCCCTTGaCTCCCAACTAGTCATAGTGTCCAGCCCTTCTTGACAGCGGAGTGTc	387
QY	291	ccacagccgagctccacagcaaccaactacatctactaacctatgaccaggtgaagacagcca	350
Db	386	CCACAGCCGGCTCCAGACACCACtCACACtATCTTAACCATtTGACCAGStGANGACCA	327
QY	351	tgtggaatgaactactaataagaaacccctagctctgacctgaaccacccgcccgaagcggagc	410
Db	326	TGTGAGAAATGATACATAGACAACCTTAGCTGGCCCTGACCAACGGGCCCAAGGGAGC	267
QY	411	ccggaggcgggcccccaagactgtgccccgaagcccgcccgctgtgaccaggatgtcaacca	470
Db	266	CCGGGGGGGGCCCCAAGAGCTGGGGCCCGAGCCCGCCGCTGTGACCAGGATGTCAACCA	207
QY	471	ccattgatactccttaagcgggcccccaactctgtgtagagaaagaaagacatccttc	530

Accession	Sequence	Position
D8	CGATTGATTCCTTACAGGGGGCCCCAAGTCTGTGAGAGAGAGAGAGACATC	- C 149
D7	531 tgaccacagcgcctcttagaccacatgtagcaacacacccgtgagctgtagccaggctccacaag	590
D6	148 TGACCACAGGGCTCTTAGACCCACATGAGCACCCACCCGCTGGCTGACAGGAGCTTCACCAAG	89
D5	591 ggcctgtgacatccagcccaagatggtccaaatgcaacgcgtctgacctaaagggccggc	650
D4	88 GGCCTGTGGCATCCAGGCCCAAGGTGTCACATGTGACACCCCTGGACCTCAAGGGCCCGGC	29
D3	651 ggtccacacccagcttgacacaagcaactc 678	
D2	28 GGTCCACGCGAGCTGGCCAAAGCACTAC 1	

Search completed: May 11, 2002, 06:19:13
Job time: 14498 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2002, 21:32:29 ; Search time 1500.46 Seconds

(without alignments)
14950.042 Million cell updates/sec

Title: US-10-082-902-1

Perfect score: 1662
Sequence: 1 aggtacgagctgagctgac.....tgctctgggaggtgtgtaga 1662

Scoring table: IDENTITY_NUC

Gapop 10.0 ; Gapext 1.0
Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estchum:*
3: em_estin:*
4: em_estnu:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_hic:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	498.2	30.0	679	9	BB648007 BB648007
2	475.4	28.6	666	9	BB630302 BB630302
3	383.6	23.1	448	10	BE709917 MI-P-VY1-BE709917
4	373	22.4	471	10	BE721232 188336 MA-BE721232
5	308.6	18.6	456	9	AA499432 AA499432
6	308	18.5	726	10	BE579947 BE579947
7	279.8	16.8	377	9	A1643238 A1643238
8	233.2	14.0	756	12	CNS023FM CNS023FM
9	222.6	13.4	557	9	AM128750 AM128750
10	202.2	12.2	401	11	AK014578 AK014578
11	201.8	12.1	401	9	A1545888 A1545888
12	201.8	11.7	657	12	AZ31672 AZ31672
13	195	11.3	1160	12	CNS03FKP CNS03FKP
14	188.6	11.3	1023	10	BE580057 BE580057
15	186.8	11.2	802	10	BI914916 BI914916
16	186	11.2	562	10	BJ032161 BJ032161
17	181	10.9	584	9	AW959446 AW959446

18	180.2	10.8	527	10	BE270351 BE270351
19	180.2	10.8	666	10	BE296059 BE296059
20	180.2	10.8	1719	11	BC004205 BC004205
21	179.4	10.8	439	9	A1613904 A1613904
22	178.4	10.7	522	9	AW463200 AW463200
23	178.2	10.7	517	9	AM266965 AM266965
24	178.2	10.7	522	10	BE040166 BE040166
25	178.2	10.7	522	10	BE044187 BE044187
26	178.2	10.7	705	9	AV650436 AV650436
27	177	10.6	531	10	BE295152 BE295152
28	172.4	10.4	720	9	AL652823 AL652823
29	172.4	10.4	720	10	BG116037 BG116037
30	172.2	10.4	823	10	BF122111 BF122111
31	169.8	10.2	609	9	AA431912 AA431912
32	167.6	10.1	976	12	CNS040JO CNS040JO
33	164	9.9	710	10	BE294735 BE294735
34	160.4	9.7	688	10	BJ060732 BJ060732
35	150.6	9.1	552	9	AW955843 AW955843
36	146.4	8.8	493	10	BF090636 BF090636
37	145.4	8.7	958	9	AL534238 AL534238
38	140.6	8.5	542	9	BE030374 BE030374
39	137.6	8.3	388	10	BE934390 BE934390
40	137.6	8.3	732	10	BI152022 BI152022
41	135.4	8.1	368	9	AA367247 AA367247
42	135.4	8.1	421	10	R69044 R69044
43	135.2	8.1	628	10	BE482909 BE482909
44	131.6	7.9	609	10	BE892466 BE892466
45	129.4	7.8	402	10	BF933262 BF933262

ALIGNMENTS

RESULT 1
LOCUS BB648007
DEFINITION BB648007 RIKEN full-length enriched, 16 days embryo head Mus
ACCESSION BB648007
VERSION BB648007.1 GI:16482263
KEYWORDS EST.
SOURCE house mouse.
ORANISM Mus musculus.
REFERENCE 1 (bases 1 to 679)
Arakawa,T., Carrincci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,T., Ishii,Y., Ito,M., Kawai,J., Komoto,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takehashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carrincci,P., Shibata,Y., Hayata,N., Sugahara,Y., Shibata,K., Itoh,M., Komoto,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matsuhira,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.

[illegible]

Db	181	CATGTGTCGTGCCTTGTCTGTCTGTACTCGCCGCCACAGCAGCTGGCAAGCTGGCCC	240
QY	995	aagcgtggtctaagaacgcgtctgcgcg-cccttgattccgcttcgaagcaacagac-aggct	1052
Db	241	AACGAGAGCTACGACCCTTTTGTCAGACGCCCTCCGGTTGGCTGTGAAGCACGAAACAAGCCT	300
QY	1053	cactctcaagaagcagcagcggggatgccaaagcacagcagcccgacaagaccttctgaca	1112
Db	301	CATTCTGCAAGGAGNCACGTGGGGACACACCAAGACACGAGTTTGACAGCCTTTCTGACA	360
QY	1113	gtttgtgtcgaag	1124
Db	361	CTTTGGATGAAA	372
RESULT	8		
CNS023FM/C	LOCUS		
DEFINITION	CNS023FM	756 bp	DNA
DEFINITION	Tetraodon nigriviridis genome survey sequence T7 end of clone		Linear
DEFINITION	23B14 of library G from Tetraodon nigriviridis, genomic survey		
DEFINITION	sequence.		
ACCESSION	AL179419.1	GI:7817476	
VERSION	GSS: genome survey sequence.		
KEYWORDS	Tetraodon nigriviridis.		
SOURCE	Tetraodon nigriviridis.		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;		
	Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;		
	Tetraodontidae; Tetraodon.		
REFERENCE	1 (bases 1 to 756)		
AUTHORS	Roeest-Crollius,H., Jalllon,O., Dasilva,C., Fitzames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.		
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigriviridis		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 756)		
AUTHORS	Roeest-Crollius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fitzames,C., Wincker,P., Brotlier,P., Quetier,F., Saurin,W. and Weissenbach,J.		
TITLE	Human gene number estimate provided by genome wide analysis using Tetraodon nigriviridis DNA sequence		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 756)		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-APR-2000) to the EMBL/Genbank/DBJ databases		
COMMENT	This sequence is a single read and was generated as part of a large scale clone-and-sequencing project of the Tetraodon nigriviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon .		
FEATURES			
source	Location/Qualifiers		
	1..756		
	/organism="Tetraodon nigriviridis"		
	/db_xref="taxon:99883"		
	/clone="23B14"		
	/clone_1lb="G"		
	/note="Genoscope sequence ID : COAG232DA07LP1-end : T7"		
BASE COUNT	133 a	210 c	265 g
ORIGIN		142 t	6 others
Query Match	Best Local Similarity	14.0%:	Score 233.2; DB 12; Length 756;
Matches	308; Conservative	1; Mismatches	105; Indels 3; Gaps 1;
QY	654	ccccaccagagtggacaagcactctcttcctgtagcagcgccgttggaagttaaatgcaa	713
Db	527	CTCCCCGGCCCAAGAAGACACCTCTCTCTGCGAAGAACGTCCGANAAGTCCCATGAC	468
QY	714	ggagatgtcatcccccggacgttgtcctctcgtctggtctgtaaccaagagatgcctgtg	773

KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
EST. zebrafish.	Danio rerio	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.	1 (bases 1 to 401)	Clark, M., Johnson, S. L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, R., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swallier, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Rittler, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.	Mashu Zebrafish EST project 1998	Unpublished (1998)	Contact: Stephen L. Johnson Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: zbrfish@wustl.wustl.edu
							CDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by: Matthew Clark. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Genome Systems, St. Louis, Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and RessourcenzentrumPrimaatDatenbank, Berlin, Germany (web address: www.rpzd.de)
							Seq primer: T3 ET from Amersham High quality center stop: 387 POLYA-No.


```
Db 287 CACCAGACCAAGTCCCTGGTCAATAGTCTGAAGGGCAATCCGACCCAGCCCAAGCAAC 346
QY 617 tccactgccaagccgttgagacctcaaggcccgagctcccaaccgagcttgacaagcaact 676
Db 347 T-----GATTGTGATGATGACTGAAGGGTCTCCTTAAAGAGGACCTGACACAGCACA 397
QY 677 tcttgcttgccaagccgttggaagtgtaaatgcaaggagtgtagcaltcccccggaagt 736
Db 398 AGTTCATTGTGTAACAGTGTGGAGTGCAAAGTGGAGAAATGCACCTGCTCCAGGAGACC 457
QY 737 tgccttccttgctgggttcgcaacaggagtgctgtgtctcagccagactctgtcaact 796
Db 458 TACCATCTCTTTGGCTGTGAACGGCAGTGCTTGTGCTGTGAGAGCATGGTGGAAAT 517
QY 797 atggaacgtgcaatgtgttggtgcaaggacatcttaccacttgcaagaaatgagcaatg 856
Db 518 ATGGAACCTGCATGTGTAGTCAAGGGCATCTTACCACCTGCTCCCATATGACGAGAG 577
QY 857 agggctctgctgagcaacacccctgctcctgctcccgctccaactgtgtgcccgtgtgt 916
Db 578 GGGATTCTCTATTCAAGATATATCTTCTGCTGTCAATCACTGCTGCTTAGATACC 637
QY 917 ccttcattgggtgctctctcgtggtgtgctgacctgctgtctgtacctgtgcaacgt 976
Db 638 TGTGATGGAGCCCATGTCTTATTATTTTTACCTTGCTTAGTGTATCTCTGCTAAG 697
QY 977 gctgcgtgaaagcttgcccaagcttgctacagacgtctgagccgcccgtgtgcccgtgca 1036
Db 698 GATGCGCTGAGAGCTGTGAGAGGTGTATGACTGATGCATCGCCAGAGGTGCAATGTA 757
QY 1037 agcaacagaaacagcgtcaltcgtcaagcag 1066
Db 758 AGAACTCCAAACATGTCTATTGTAGCTG 787
```

Search completed: May 11, 2002, 02:42:24
Job time: 18395 sec

QY	303	ccagaccacactcaacactccctaccacatttgaacaggttgaaagaccacacactgtgaggaatga	362
Db	337	ccagaccacactcaacactccctaccacatttgaacaggttgaaagaccacacactgtgaggaatga	396
QY	363	ctaacataagacaacaccttgaagctctgagcccttgagacacacccgccccaaagcgagacaacccgagcgagc	422
Db	397	ctaacataagacaacaccttgaagctctgagcccttgagacacacccgccccaaagcgagacaacccgagcgagc	456
QY	423	ccccagagcttgagcccccagacgcccgcgcgtgtgaacagatgttcaacccacttgatctc	482
Db	457	ccccagagcttgagcccccagacgcccgcgcgtgtgaacagatgttcaacccacttgatctc	516
QY	483	cttcaagagggagcccccaagcctctgtgagcagcagacagacatactctcttgaccaagcgt	542
Db	517	cttcaagagggagcccccaagcctctgtgagcagcagacagacatactctcttgaccaagcgt	576
QY	543	cttgaagacacatactgacacacacaccccttgagcttgacacagagccttcaacaaaggcgtgtgacat	602
Db	577	cttgaagacacatactgacacacacaccccttgagcttgacacagagccttcaacaaaggcgtgtgacat	636
QY	603	ccagaccacaagcttgagtcacacttgacacgcgcgttgagacctcaaaagggcccgagcttccaccga	662
Db	637	ccagaccacaagcttgagtcacacttgacacgcgcgttgagacctcaaaagggcccgagcttccaccga	696
QY	663	gctgtgacaagaagcactctctgtgtgtgcagagccgttgaggaaagttaaatgtcaagagatgtgc	722
Db	697	gctgtgacaagaagcactctctgtgtgtgcagagccgttgaggaaagttaaatgtcaagagatgtgc	756
QY	723	atccccccagagacgtttgacctctccgcgcgtggagctctgacaaacagagatgacgttgatgtctgaagcca	782
Db	757	atccccccagagacgtttgacctctccgcgcgtggagctctgacaaacagagatgacgttgatgtctgaagcca	816
QY	783	gactctgagtcacaactatgagcagctgtgacatgtgtgtgtgcagaggtgacatctctcaacacttgac	842
Db	817	gactctgagtcacaactatgagcagctgtgacatgtgtgtgtgtgcagaggtgacatctctcaacacttgac	876
QY	843	gaatgaaagaaagaaatgaaagggctccctgcgcgtgtgacacacccctctctgcctccgcgtcccaactgt	902
Db	877	gaatgaaagaaagaaatgaaagggctccctgcgcgtgtgacacacccctctctgcctccgcgtcccaactgt	936
QY	903	ctgagcccccgcgtgagcttcaatgttgagctctctccctccgcgtgagctgcgcgcctgacgtctgtga	962
Db	937	ctgagcccccgcgtgagcttcaatgttgagctctctccctccgcgtgagctgcgcgcctgacgtctgtga	996
QY	963	ccctgctctgacacacccgcgtgcgttgaaagcttgagccacgcgttgagcttgaacccgtctctgcgcgcgc	1022
Db	997	ccctgctctgacacacccgcgtgcgttgaaagcttgagccacgcgttgagcttgaacccgtctctgcgcgcgc	1056
QY	1023	tggtttgcgcgcgtgtgacaagacacagaacagcgttcaatctgtcaaaagcagccagccgaggtgagtcacaa	1082
Db	1057	tggtttgcgcgcgtgtgacaagacacagaacagcgttcaatctgtcaaaagcagccagccgaggtgagtcacaa	1116
QY	1083	gaccagagcagagccccgaacaaagccttctctgagacagtttgtgtctgaagccccagtgctctgcgcgt	1142
Db	1117	gaccagagcagagccccgaacaaagccttctctgagacagtttgtgtctgaagccccagtgctctgcgcgt	1176
QY	1143	gaaacccgtgctctctctctgacacatctaaagaaagacatgacagacaagtgacaagagcttcttgacctc	1202
Db	1177	gaaacccgtgctctctctctgacacatctaaagaaagacatgacagacaagtgacaagagcttcttgacctc	1236
QY	1203	ctgagagctgtgaccttgctagtctgtgcacactccctacccccagagcttcgnaaaatataagagac	1262
Db	1237	ctgagagctgtgaccttgctagtctgtgcacactccctacccccagagcttcgnaaaatataagagac	1296
QY	1263	cacacacacacgtacacctgtatctccccaagatagtgaagaaagacatttgaggctcttttttca	1322
Db	1297	cacacacacacgtacacctgtatctccccaagatagtgaagaaagacatttgaggctcttttttca	1356
QY	1323	gggtctcttaaaacttgctgtctcaaaacagacaaatgacagggcagaggtgtgagtttgaggggagaa	1382
Db	1357	gggtctcttaaaacttgctgtctcaaaacagacaaatgacagggcagaggtgtgagtttgaggggagaa	1416

```

QY 1383 tttctctttcgaagacaagaacatgttgacacatatccggaactgcagctgct 1442
Db 1417 tttctctttcgaagacaagaacacgaatgttgacacatatccggaactgcagctgct 1476
QY 1443 tgaatgcctccacagccctctctccctccctccgccccccctctctcttc 1502
Db 1477 tgaatgcctccacagccctctctctctccctccctccgccccccctctcttc 1536
QY 1503 cattgtctttggacaccaaagagctagctgcctcggagaggaattgcctaactgtacag 1562
Db 1537 c-ttgtctttggacaccaaagagctagctgcctcggagaggaattgttaactgtacag 1595
QY 1563 gtaaccttaaaagaacaccttgagagctctctacatacctctctccctccctccatcact 1622
Db 1596 gtaaccttaaaagaacaccttgagagctctctacatacctctctccctccatcact 1655
QY 1623 ccaccaccattgtccctgaatgtcttggaggaaagtgtaga 1662
Db 1656 ccaccaccattgtccctgaatgtcttggaggaaagtgtaga 1695

RESULT 3
PCT-US02-09921-510
; Sequence 510, Application PC/US0209921
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: DAFPO, Abel
; APPLICANT: JONES, Anissa L.
; APPLICANT: TRAN, Alanna-Phung B.
; APPLICANT: DAHL, Christopher R.
; APPLICANT: GIETZEN, Darryl
; APPLICANT: CHINN, Joyce
; APPLICANT: DUFOUR, Gerard E.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: YU, Jimmy Y.
; APPLICANT: TUASON, Olivia
; APPLICANT: YAP, Pierre E.
; APPLICANT: AMSHEY, Stefan R.
; APPLICANT: DAUGHERTY, Sean C.
; APPLICANT: DAM, Tam C.
; APPLICANT: LIU, Tommy F.
; APPLICANT: NGUYEN, Duy-Viet An
; APPLICANT: KLEEFELD, Yaël
; APPLICANT: GERSTIN JR., Edward H.
; APPLICANT: PERALTA, Careyana H.
; APPLICANT: DAVID, Marie A.
; APPLICANT: LEWIS, Samantha A.
; APPLICANT: CHEN, Alice J.
; APPLICANT: PANZER, Scott R.
; APPLICANT: HARRIS, Bernard
; APPLICANT: FLORES, Vincent
; APPLICANT: MARAÑA, Rakesh
; APPLICANT: LO, Audrey
; APPLICANT: IAN, Ruth Y.
; APPLICANT: URASHKA, Michael
; TITLE OF INVENTION: SECRETORY MOLECULES
; FILE REFERENCE: PT-1232 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/09921
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 60/280, 067; 60/280, 068; 60/291, 280; 60/291, 849;
60/291, 829; 60/299, 428; 60/300, 001; 60/299, 776
; PRIOR FILING DATE: 2001-03-29; 2001-03-29; 2001-05-16; 2001-05-17;
2001-05-17; 2001-06-19; 2001-06-20; 2001-06-20
; NUMBER OF SEQ. ID NOS: 1146
; SOFTWARE: PERL program
; SEQ ID NO 510
; LENGTH: 5698
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: LG:241742.1:2001JUN22

```

OY	1164	atctaaagaagctctcagaacagatgcacaaaggttttttagctctctgaagcttgaaccttgtaac	1223
Dp	1976	atctaaagaagctctcagaacagatgcacaaaggttttttagctctctgaagcttgaaccttgtaac	2035
OY	1224	tgcccaactccctaacccccagctcttcgnaaaatacagaagaccacacacacgtaacctgatac	1283
Dp	2036	tgcccaactccctaacccccagctctcttcgnaaaatacagaagaccacacacacgtaacctgatac	2095
OY	1284	ccccaaagatgatctgaagaagaagacctcttgaggctttttctcaagggtccctgaaccttgatca	1343
Dp	2096	ccccaaagatgatctgaagaagaagacctcttgaggctttttctcaagggtccctgaaccttgatca	2155
OY	1344	aacagacaatgacaggggcagggctgctggcttgggggggnaaaattttcttttcagaagaacag	1403
Dp	2156	aacagacaatgacaggggcagggctgctggcttgggggggnaaaattttcttttcagaagaacag	2215
OY	1404	aacacagaatctggagacacatactccggaagaacctcgacgcgtcttgatgctctccagaccctc	1463
Dp	2216	aacacagaatctggagacacatactccggaagaacctcgacgcgtcttgatgctctccagaccctc	2275
OY	1464	ctctccctccctccctccctcccgccccccctctctctcttcacatgtgctcttgacatccacag	1523
Dp	2276	ctctccctccctccctccctcccgccccccctctctctcttcacatgtgctcttgacatccacag	2335
OY	1524	gagcttagctctgcctgggagggagaaattgcttaactgagttacacagggtaacctttaagaagaacct	1583
Dp	2336	gagcttagctctgcctgggagggagaaattgcttaactgagttacacagggtaacctttaagaagaacct	2395
OY	1584	tggagctctctataactctctctccctctccacatccatccacacacacacacacttgcctgat	1643
Dp	2396	tggagctctctataactctctctccctctccacatccatccacacacacacacacttgcctgat	2455
OY	1644	gtctctgggagaggtgtataga 1662	
Dp	2456	gtctctgggagaggtgtataga 2474	

Query Match	13.9%	Score 230.2;	DB 6;	Length 3475;
Best Local Similarity	98.7%;	Pred. No. 8.1e-43;		
Matches 232; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

1488 cccctctcttttccatgtctcttcttgcaactcaacagagactgcgcgcctgtagaagaattg 1547

```

Db      61 cccctcccttttccatctgtctgtgtctctccacagagctagctgccttgaggagattg 120
        |||
Qy      1548 ctactgagtlacagagggtactcttaagaagaacccttgagctctccatccctctc 1607
        |||
Db      121 ttaactgagtlacagagggtactcttaagaagaacccttgagctctccatccctctc 180
        |||
Qy      1608 ctcccatctactccacccacacttgctccctgagctgcttgaggaggtgtaga 1662
        |||
Db      181 ctcccatctactccacccacacttgctccctgagctgcttgaggaggtgtaga 235
        |||

```

RESULT 5

```

US-10-105-299-8306
; Sequence 8306, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950
; CURRENT APPLICATION NUMBER: US/10/105,299
; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8306
; LENGTH: 2227
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-105-299-8306

```

```

Query Match      10.8%; Score 180.2; DB 6; Length 2227;
Best Local Similarity 66.7%; Pred. NO. 1.2e-31;
Matches 273; Conservative 0; Mismatches 133; Indels 3; Gaps 1;

```

```

Qy      661 gagctgagacaagcactctgtctgtgagagcctgtggaagtgttaattgcaagaagtgt 720
        |||
Db      1005 gatttggcctcagcagcctacagagtgtagagagctgagcaagtgcataagagagtgct 1064
        |||
Qy      721 gcatcccccggagcgtgtgctctccctgctgctgctgctgctgctgctgctgctgctg 780
        |||
Db      1065 acctaccacagagcctctgcatcagactgagctgagcaagaagcgtcttgcctgcgc 1124
        |||
Qy      781 cagactctgtcaacatgacagctgacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 840
        |||
Db      1125 cagaagctgattgactatgagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1184
        |||
Qy      841 acgaatgagagcagtgagcgtctgctgctgctgctgctgctgctgctgctgctgctgctg 900
        |||
Db      1185 tctaat---gatgagagagacaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1241
        |||
Qy      901 tgcgtgccccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
        |||
Db      1242 tgtgtacacagatgagcagcagctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1301
        |||
Qy      961 tacctgtctgacacagcgtgctgagagctgagcagcgtgtgtgtgtgtgtgtgtgtgtgt 1020
        |||
Db      1302 taccttccagcagaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1361
        |||
Qy      1021 cctgtgtgctgctgcaagcagcagcagcagcagcagcagcagcagcagcagcagcag 1069
        |||
Db      1362 cctgtgtgctgctgcaagcagcagcagcagcagcagcagcagcagcagcagcagcag 1410
        |||

```

```

RESULT 6
US-10-105-299-8308
; Sequence 8308, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950
; CURRENT APPLICATION NUMBER: US/10/105,299
; CURRENT FILING DATE: 2002-03-26

```

```

; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8308
; LENGTH: 2227
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-105-299-8308

```

```

Query Match      10.8%; Score 180.2; DB 6; Length 2227;
Best Local Similarity 66.7%; Pred. NO. 1.2e-31;
Matches 273; Conservative 0; Mismatches 133; Indels 3; Gaps 1;

```

```

Qy      661 gagctgagacaagcactctgtctgtgagagcctgtggaagtgttaattgcaagaagtgt 720
        |||
Db      1005 gatttggcctcagcagcctacagagtgtagagagctgagcaagtgcataagagagtgct 1064
        |||
Qy      721 gcatcccccggagcgtgtgctctccctgctgctgctgctgctgctgctgctgctgctg 780
        |||
Db      1065 acctaccacagagcctctgcatcagactgagctgagcaagaagcgtcttgcctgcgc 1124
        |||
Qy      781 cagactctgtcaacatgacagctgacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 840
        |||
Db      1125 cagaagctgattgactatgagcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1184
        |||
Qy      841 acgaatgagagcagtgagcgtctgctgctgctgctgctgctgctgctgctgctgctgctg 900
        |||
Db      1185 tctaat---gatgagagagacaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1241
        |||
Qy      901 tgcgtgccccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
        |||
Db      1242 tgtgtacacagatgagcagcagctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1301
        |||
Qy      961 tacctgtctgacacagcgtgctgagagctgagcagcgtgtgtgtgtgtgtgtgtgtgtgt 1020
        |||
Db      1302 taccttccagcagaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1361
        |||
Qy      1021 cctgtgtgctgctgcaagcagcagcagcagcagcagcagcagcagcagcagcagcag 1069
        |||
Db      1362 cctgtgtgctgctgcaagcagcagcagcagcagcagcagcagcagcagcagcagcag 1410
        |||

```

```

RESULT 7
US-10-105-299-7769/c
; Sequence 7769, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950
; CURRENT APPLICATION NUMBER: US/10/105,299
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7769
; LENGTH: 4802
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-105-299-7769

```

```

Query Match      10.8%; Score 180.2; DB 6; Length 4802;
Best Local Similarity 66.7%; Pred. NO. 1.4e-31;
Matches 273; Conservative 0; Mismatches 133; Indels 3; Gaps 1;

```

```

Qy      661 gagctgagacaagcactctgtctgtgagagcctgtggaagtgttaattgcaagaagtgt 720
        |||
Db      1223 gatttggcctcagcagcctacagagtgtagagagctgagcaagtgcataagagagtgct 1164
        |||
Qy      721 gcatcccccggagcgtgtgctctccctgctgctgctgctgctgctgctgctgctgctg 780
        |||
Db      1163 acctaccacagagcctctgcatcagactgagctgagcaagaagcgtcttgcctgcgc 1104
        |||

```

```

QY 781 cagaactggtcaactatgacgtgcatgtgttgcgaaggacatctctacactgc 840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1103 CAGAACGTGATTGACTATGGACTTGTGTATGCTGTGAAGGTCCTTCTATCACTGT 1044
QY 841 acgaatgagacgtatgaggtctctgctctacacccctgctctctccgtccac 900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1043 TCTAAT---GATGATGAGGACACTGTGTCACAAACCATGTTCTTGCAGCACTGCAC 987
QY 901 tgcctgcccgcgtggtctctctacatgaggtctctctgaggtgctgcccgtctctgc 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 986 TGTGTACACAGATGTCAGCATGGTGTGTCATGTCCTCTTTTGGCTTGTATGCTGT 927
QY 961 taactgctgacccgctgctgctgagagctgcccagcgtgctacacgctctgcgcgc 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 926 TACCTTCAGCCAGAGGCTGCTTAATTTGTCAGGAGGCTGTATGACGGGCTTACAGC 867
QY 1021 cctggttcgcgtcgaagcaacgaagcgtcatctctgcaagaagcaca 1069
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 866 CCGTGTGCCGCTGTAAAACTCAACACAGTTTGTGCAAGATTCCCA 818

```

```

RESULT 8
: Sequence 8307, Application US/10105299
: GENERAL INFORMATION:
: APPLICANT: Rosen, et. al
: TITLE OF INVENTION: Human Secreted Proteins
: FILE REFERENCE: PS950
: CURRENT APPLICATION NUMBER: US/10/105,299
: NUMBER OF SEQ ID NOS: 15197
: Prior Application removed - See File Wrapper or Palm
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 8307
: LENGTH: 4802
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-105-299-8307

```

```

Query Match 10.8%; Score 180.2; DB 6; Length 4802;
Best Local Similarity 66.7%; Pred. No. 1.4e-31;
Matches 273; Conservative 0; Mismatches 133; Indels 3; Gaps 1;

QY 661 gagctggaacagcactctctctgctgagcgcctgctggaagctgaatgcaagagtg 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3580 gatttggtccctgacgctacaggtgtgagagctgtgcaagtgcaaatgtaagagtg 3639
QY 721 gaatcccccgaagctgctctctctctgctgagcgtgcaagagctgctgctcagcc 780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3640 acctaccacaagcctctgctacatgactgctgctgcaagcagctgctctgctgcgc 3639
QY 781 cagactctgtcaactatgacgtgcatgtgtgtgctgagggacatctctacacactgc 840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3700 cagaacgtgattgactatgagactgtgtatgtgtgtgtgaaagctctctctatcaactgt 3759
QY 841 acgaatgagacgtatgaggtctctgctgctgacacccctgctctctccgtccac 900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3760 tctaatt---galtgtagagacacactgtctgcaacccatctctctgcaagcagctac 3816
QY 901 tgcctgcccgcgtggtctctctacatgaggtctctctgaggtgctgcccgtgctgcgc 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3817 tgtgttaacagatgctgacagcagtggtgtctatgtctccctcttctgctgtttaaagtgt 3876
QY 961 taactgctgacccgctgctgctgagagctgcccagcgttggtctacagcgtctgcgcgc 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3877 taacttcagacgaaggtgtgctcttaattgtgctcaggggtgttatgacgggtttaacag 3936
QY 1021 cctggttcgcgtcgaagcaacgaagcgtcatctctgcaagaagcaca 1069
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3937 cctggttcgcgtcgaagcaacgaagcgtcatctctgcaagaagcaca 3985

```

```

RESULT 9
US-10-105-299-9313/c
: Sequence 9313, Application US/10105299
: GENERAL INFORMATION:
: APPLICANT: Rosen, et. al
: TITLE OF INVENTION: Human Secreted Proteins
: FILE REFERENCE: PS950
: CURRENT APPLICATION NUMBER: US/10/105,299
: NUMBER OF SEQ ID NOS: 15197
: Prior Application removed - See File Wrapper or Palm
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 9313
: LENGTH: 4802
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-105-299-9313

```

```

Query Match 10.8%; Score 180.2; DB 6; Length 4802;
Best Local Similarity 66.7%; Pred. No. 1.4e-31;
Matches 273; Conservative 0; Mismatches 133; Indels 3; Gaps 1;

QY 661 gagctggaacagcactctctctgctgagcgcctgctggaagctgaatgcaagagtg 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1223 GATTGGGCTGACAGCGCTACAGGTGTGAGACTGTGCAAGTGCMAAATGTAAGAGTGC 1164
QY 721 gaatcccccgaagctgctctctctctgctgagcgtgcaagagctgctgctcagcc 780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1163 ACTTACCCCAAGGCTCTGTCATCAGACTGCATGTGTCGACAAAGTGCCTTCTGCGCC 1104
QY 781 cagactctgtcaactatgacgtgcatgtgtgtgctgagggacatctctacacactgc 840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1103 CAGAACGTGATTGACTATGGACTTGTGTATGCTGTGAAGGTCCTTCTATCACTGT 1044
QY 841 acgaatgagacgtatgaggtctctgctgctgacacccctgctctctccgtccac 900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1043 TCTAAT---GATGATGAGGACACTGTGTCACAAACCATGTTCTTGCAGCACTGCAC 987
QY 901 tgcctgcccgcgtggtctctctacatgaggtctctctgaggtgctgcccgtctctgc 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 986 TGTGTACACAGATGTCAGCATGGTGTGTCATGTCCTCTTTTGGCTTGTATGCTGT 927
QY 961 taactgctgacccgctgctgctgagagctgcccagcgttggtctacagcgtctgcgcgc 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 926 TACCTTCAGCCAGAGGCTGCTTAATTTGTCAGGAGGCTGTATGACGGGCTTACAGC 867
QY 1021 cctggttcgcgtcgaagcaacgaagcgtcatctctgcaagaagcaca 1069
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 866 CCGTGTGCCGCTGTAAAACTCAACACAGTTTGTGCAAGATTCCCA 818

```

```

RESULT 10
US-10-105-299-11466/c
: Sequence 11466, Application US/10105299
: GENERAL INFORMATION:
: APPLICANT: Rosen, et. al
: TITLE OF INVENTION: Human Secreted Proteins
: FILE REFERENCE: PS950
: CURRENT APPLICATION NUMBER: US/10/105,299
: NUMBER OF SEQ ID NOS: 2002-03-26
: Prior Application removed - See File Wrapper or Palm
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 11466
: LENGTH: 4802
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-105-299-11466

```

```

Query Match 10.8%; Score 180.2; DB 6; Length 4802;

```


Page 8

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:45:48 ; Search time 16.06 Seconds

(without alignments)
1418.194 Million cell updates/sec

Title: US-10-082-902-2

Sequence: 1 MEPPIQSAPLTPNSVMWP.....SVICKAASGDAKSPDKRF 299

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: 1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	155	9.4	2	JC5909
2	120	7.2	2	T34022
3	116	7.0	2	T18355
4	113	6.8	2	A57283
5	110.5	6.7	1	VMHU
6	109	6.6	2	T14764
7	108	6.5	2	A35672
8	107.5	6.5	2	S08167
9	107	6.5	2	T25471
10	107	6.5	2	A38261
11	106.5	6.4	2	T23681
12	105	6.3	2	T42215
13	104	6.3	2	A29072
14	104	6.3	2	T27283
15	102.5	6.2	2	B37057
16	102.5	6.2	1	A53102
17	101	6.1	2	A55624
18	101	6.1	1	O0BR8
19	101	6.1	2	T34513
20	100.5	6.1	2	S04847
21	100.5	6.1	2	A45839
22	100	6.0	2	T09059
23	99.5	6.0	2	P05053
24	99	6.0	2	A49175
25	98.5	5.9	2	T15099
26	98.5	5.9	2	B26637
27	98	5.9	2	T34446
28	98	5.9	2	A40043
29	97.5	5.9	2	D86446

30	97.5	5.9	614	2	A32608	thyroid hormone re
31	97.5	5.9	798	2	A40526	integrin beta-7 ch
32	97	5.9	1113	2	JE0315	low-density lipopr
33	96.5	5.8	269	2	T40371	probable zinc fing
34	96.5	5.8	3002	2	A47221	fibritillin 1 precu
35	96	5.8	457	2	T24962	odr-7 protein - Ca
36	96	5.8	965	2	S62935	hypothetical prote
37	96	5.8	1798	2	S53869	laminin beta-2 cha
38	96	5.8	2195	2	T34264	hypothetical prote
39	95.5	5.8	772	2	S32659	integrin beta 2 ch
40	95.5	5.8	972	2	A30363	glycoprotein gp330
41	95.5	5.8	1776	1	RRAPYM	genome polyprotein
42	95.5	5.8	2531	2	S18188	notch protein homo
43	95	5.7	429	2	T21113	hypothetical prote
44	95	5.7	769	1	IJH0LM	leukocyte adhesio
45	95	5.7	769	1	JC1121	leukocyte adhesio

ALIGNMENTS

RESULT 1

JC5909 AE33 protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C>Date: 18-Mar-1998 #sequence_revision 18-Mar-1998 #text_change 07-May-1999

C:Accession: JC5909

R:Demille, M.M.C.; Kimmel, B.E.; Rubin, G.M.
Gene 183, 103-108, 1996

A:Title: A Drosophila gene regulated by rough and glass shows similarity to ena and v

A:Reference number: JC5909, PMID:97149285

A:Accession: JC5909

A:Molecule type: mRNA

A:Residues: 1-425 <DEM>

C:Comment: This protein plays a role in signalling during R2 and R5 development.

C:Genetics:

A:Keywords: phosphoprotein

F:262/Binding site: phosphate (Ser) (covalent) (by cAMP- and cGMP-dependent kinases)

F:311/Binding site: phosphate (Thr) (covalent) (by cAMP- and cGMP-dependent kinases)

Query Match 9.4% Score 155; DB 2; Length 425;

Best Local Similarity 23.3% Pred. No. 0.0003;

Matches 66; Conservative 22; Mismatches 99; Indels 96; Gaps 14;

QY 52 YIDNPSLATGPKRRGAP--ELAPTPARCDQ-----DYTHHWISF----- 92

DB 193 YISDKTSATTSP---DAPPASAPSPATACIAASEYVTTLTAVHHDTNYPVDPD 248

QY 93 -----SGRPSSVSSSSSTSSDQRLDHPMAPPVADQSPRAVRIOPKVHCOP---LD 142

DB 249 VGAQVLNARESSISALKRRNALE-----MAQMAAQTAAAGGLACDGSSGR 297

QY 143 LKGPVAPPELDKHF-LICEFC-----GKCKCKEACSPRTLPSCVNCQELCSA 190

DB 298 LHKPMVSDILKTRKRCRCYCHELYSEDEFNRGACE-----YAPDAFRSGYECIS-- 347

QY 191 QTLVNTGTCMGLVGIFFYHCTNEDGSCADHPCSGSRS--NCCARMSFGALSYVLPCL 248

DB 348 -----GMCACARMT---YHOMS-DAEETNQHPCDCSASAGSKRMGLIALISLFPCL 398

QY 249 LCVLPATGCVKLAORGVDRLRRPGCRCKHTNSVICKAASGDAK 291

DB 399 WCVPPLRAC-----HLAGIGHGLGCGGQHK 422

RESULT 2

T34022 zonadhesin - pig

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 07-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000

C:Accession: T34022

R:Hardy, D.M.; Garbers, D.L.
 J. Biol. Chem. 270, 26025-26028, 1995
 A:Title: A sperm membrane protein that binds in a species-specific manner to the egg ext
 A:Reference number: Z21464; MUID:96064638
 A:Accession: T34022
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-2476 <HAR>
 A:Cross-references: EMBL:U40024; NID:91066465; PID:91066466; PIDN:AA648486.1
 A:Experimental source: strain Welshan; testis
 A:Genetics:
 A:Gene: Zan
 A:Function:
 A:Description: may be involved in sperm adhesion to the zona pellucida

Query Match 7.28; Score 120; DB 2; Length 2476;
 Best Local Similarity 19.8%; Pred. No. 0.59;
 Matches 63; Conservative 33; Mismatches 112; Indels 110; Gaps 17;

OY 2 EPPRQSAPIIPNSVMQPLDSSHSRLQHL--TLPIDOVKTSVENDYIDNPSLA 59
 DB 552 KPYVTEPSVPTKEKPYTPTEPTKLVPTERTTPKRTTPTRIR---TTPTR 608
 OY 60 LTTGKRTGAPAPLAPPCDDVTHMTSFGSRPSSVSSSTSDQRLDHMAAPP 119
 DB 609 TTPTRIR-----TTPTRIRTPTRER--TTPTRKTPVTEKTIPTERTIR---AP 654
 OY 120 VADQASPAVRIQPKVY-----HCQPLDKGAVPELDKHLCEAC 162
 DB 655 TTPQSPPLVLPQPAVAMPSTATVTPRTTASCP-----NAHF----- 696
 OY 163 GKCKCK-SCASPRITLPSCMV-CNQCIC-----SAQLVNGCMCLVQGFHCTNED 214
 DB 697 ERGACPVSCGP--TPNCELFCRKCVCDDPGLFSGSHCVASSCDDCYNNYKRLGDM 754
 OY 215 DEGCADHPSC--SRNC-----CA----- 233
 DB 755 FSPNCTEH-CHCRPSSRMECOTFCGTHVQQLKNGQYCHPGSATCSVYGDPHYLTFD 813
 OY 234 --RMSFGALSVL--PC 247
 DB 814 GRNFNGKCTYTLAOPC 831

RESULT 3
 18355
 hypothetical protein P3 - Mycoplasma hyorhinis

C:Species: Mycoplasma hyorhinis
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 07-Dec-1999
 C:Accession: T18355
 R:Deng, G.; McIntosh, M.A.
 J. Bacteriol. 176, 5929-5937, 1994
 A:Title: An amplifiable DNA region from the Mycoplasma hyorhinis genome.
 A:Reference number: 218888; MUID:95014025
 A:Accession: T18355
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1187 <DEN>
 A:Cross-references: EMBL:L11447; NID:9150156; PID:9664956; PIDN:AAA62228.1
 A:Genetics:
 A:Genetic code: SGC3

Query Match 7.0%; Score 116; DB 2; Length 1187;
 Best Local Similarity 21.9%; Pred. No. 0.57;
 Matches 59; Conservative 25; Mismatches 113; Indels 72; Gaps 14;

OY 35 LTLIPIDOVKTSVENDYIDNPSLALTGPKRTGAPAPLAPPCDDVTHMTS 92
 DB 266 LTLDEDDHFEYVHEEY-----CLA-----CONHATCDICCNLESESLY 306
 OY 93 SGRPSSVSSSTSDQRLDHMAAPPVADQASPAVRIQPKVYHCQPLDKGAVPEL 152

DB 307 RLKNGQVNNLELELEEL-HVADSPYQKRPCCGSKLETESCD----- 352
 OY 153 DKHFLCEACGCKCKCK-----ASPRITL---PSCVNCQNELSCAQLTV-NYGCML 202
 DB 353 -----CEA-----CKQCEBNCSCSELTCGQOBATCSCAQEHGCGEESCAPNTTCACT 403
 OY 203 VQGFYHCTNEDDEGSCADHPSCSRNCCARMSFGALSVLPCLCYLPA---TGCV 258
 DB 404 EE-----HCEPTSTCGCENPCECEBACDCCSEHCEBCEVETQACLDCTNQTAKYCGCT 459
 OY 259 KLAQGYRLRLRPGC-RCKHTNSV-ICKA 285
 DB 460 -----QEQHPTCECKECDECKOCKA 480

RESULT 4

A57283
 Integrin beta chain pat-3 precursor - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 29-Oct-1999
 C:Accession: A57283; T27675
 R:Oetner, S.N.; Kenyon, C.; Reichardt, L.F.
 J. Cell Biol. 129, 1127-1141, 1995
 A:Title: Characterization of betap-3 heterodimers, a family of essential integrin
 A:Reference number: A57283; MUID:95263667
 A:Accession: A57283
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-809 <GET>
 A:Cross-references: GB:019744; NID:9644777; PIDN:AAA85704.1; PID:9644778
 R:Mortimore, B.
 submitted to the EMBL Data Library, August 1994
 A:Reference number: Z20403
 A:Accession: T27675
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-809 <ML>
 A:Cross-references: EMBL:Z25604; PIDN:CAA84677.1; GSPDB:GN00021; CESP:ZK1058.2
 A:Experimental source: clone ZK1058
 A:Genetics:
 A:Gene: CESP:ZK1058.2
 A:Map position: 3
 A:Insertions: 66/3; 135/1; 269/3; 415/2; 528/2; 691/1; 786/3
 C:Superfamily: Integrin beta chain; laminin-type EGF-like homology

Query Match 6.8%; Score 113; DB 2; Length 809;
 Best Local Similarity 26.7%; Pred. No. 0.66;
 Matches 27; Conservative 11; Mismatches 37; Indels 26; Gaps 5;

OY 162 GKCKCK-----ECASPRITLPS-----CWVNCQELCSAQLTVNYGTCMC----- 201
 DB 496 CGVRCGNGVNGKCKCECNRRGMSYALNLEKCKRNEASICEGVCNCGRCECNPRANPE 555
 OY 202 -LVQGFYHCTN-----EDDEGSCADH-PCSCSRNSNCCARMS 236
 DB 556 EQLSGEFCEDCNFNCPRHDKRICAEHGECNCGKICAPGWT 596

RESULT 5

VWU
 von Willebrand factor precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 04-Dec-1986 #sequence_revision 30-Jun-1993 #text_change 22-Jun-1999
 C:Accession: A34480; S02377; A37139; S23676; A25298; A25469; A25366; S23618; S23645;
 R:Mancuso, D.J.; Tukey, E.A.; Westfield, L.A.; Mortill, N.K.; Shelton-Inloes, B.B.; S
 J. Biol. Chem. 264, 19514-19527, 1989
 A:Title: Structure of the gene for human von Willebrand factor.
 A:Reference number: A34480; MUID:90062044
 A:Accession: A34480
 A:Molecule type: DNA
 A:Residues: 1-2813 <MAN>

A:Cross-references: EMBL:M25864
R:Bonthron, D.; Orkin, S.H.
Eur. J. Biochem. 171, 51-57, 1988
A:Title: The human von Willebrand factor gene. Structure of the 5' region.
A:Reference number: S02377; MUID:88111704
A:Accession: S02377
A:Molecule type: DNA
A:Residues: 1-177 <BO2>
A:Cross-references: EMBL:X06828
R:Mancuso, D.J.; Tulley, E.A.; Westfield, L.A.; Lester-Mancuso, T.L.; Le Beau, M.M.; Sorio
Biochemistry 30, 253-269, 1991
A:Title: Human von Willebrand factor gene and pseudogene: structural analysis and differ
A:Accession: A37139
A:Molecule type: DNA
A:Residues: 990-1947 <MAD>
A:Cross-references: GB:M60675; NID:9340357; PIDN:AAA61295.1; PID:9553810
A:Note: the authors translated the codon CGC for residue 156 as Gln
R:Collins, C.J.; Underdahl, J.P.; Levene, R.B.; Ravera, C.P.; Morin, M.J.; Dombalagian,
Proc. Natl. Acad. Sci. U.S.A. 84, 4393-4397, 1987
A:Title: Molecular cloning of the human gene for von Willebrand factor and identificatio
A:Reference number: S23676; MUID:87260814
A:Accession: S23676
A:Molecule type: DNA
A:Residues: 2731-2813 <COL>
A:Cross-references: EMBL:M16945
R:Bonthron, D.; Orr, E.C.; Mitsuoka, L.M.; Ginsburg, D.; Handin, R.I.; Orkin, S.H.
Nucleic Acids Res. 14, 7125-7127, 1986
A:Title: Nucleotide sequence of pre-pro-von Willebrand factor cDNA.
A:Reference number: A25298; MUID:87016349
A:Accession: A25298
A:Molecule type: mRNA
A:Residues: 1-470, 'V', 472-2813 <BON>
A:Cross-references: EMBL:X04385
R:Verweij, C.L.; Diergaarde, P.J.; Hart, M.; Pannekoek, H.
EMBO J. 5, 1839-1847, 1986
A:Title: Full-length von Willebrand factor (vWF) cDNA encodes a highly repetitive protei
A:Reference number: A91044; MUID:87004550
A:Accession: A25469
A:Molecule type: mRNA
A:Residues: 1-470, 'V', 472-483, 'R', 485-1022, 'K', 1024-1025, 'E', 1027-1400 <VER>
A:Cross-references: EMBL:X04146
A:Note: this sequence has been revised in reference A91056
R:Verweij, C.L.; Diergaarde, P.J.; Hart, M.; Pannekoek, H.
EMBO J. 5, 3074, 1986
A:Reference number: A91056
A:Accession: A25366
A:Molecule type: mRNA
A:Residues: 1021-1030 <VE2>
A:Note: this is a revision to the sequence from reference A91044
R:Shelton-Inloes, B.B.; Broeze Jr., G.J.; Miletich, J.P.; Sadler, J.E.
Biochem. Biophys. Res. Commun. 144, 657-665, 1987
A:Title: Evolution of human von Willebrand factor: cDNA sequence polymorphisms, repeated
A:Reference number: S23618; MUID:87213253
A:Accession: S23618
A:Molecule type: mRNA
A:Residues: 1-120 <SH2>
A:Cross-references: EMBL:M17588; NID:9799330; PIDN:AAA65940.1; PID:9340316
A:Accession: S23645
A:Molecule type: protein
A:Residues: 23-56 <SH3>
R:Sadler, J.E.; Shelton-Inloes, B.B.; Sorace, J.M.; Harlan, J.M.; Titani, K.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 82, 6394-6398, 1985
A:Title: Cloning and characterization of two cDNAs coding for human von Willebrand facto
A:Reference number: A94060; MUID:86016708
A:Accession: A94060
A:Molecule type: mRNA
A:Residues: 'WA', 739, 'C', 744-769, 'H', 771-788, 'A', 790-803, 'S', 805-873, 1289-1471, 'D', 1473-
A:Note: the authors translated the codon TGC for residue 2166 as Cys
R:Shelton-Inloes, B.B.; Titani, K.; Sadler, J.E.
Biochemistry 25, 3164-3171, 1986
A:Title: cDNA sequences for human von Willebrand factor reveal five types of repeated do
A:Reference number: A90504; MUID:86269894

A:Accession: A90504
A:Molecule type: mRNA
A:Residues: 781-788 'A', 790-1424 <SHE>
A:Note: 852-Gln, 857-Asp, and 1381-Thr were also found
R:Ginsburg, D.; Handin, R.I.; Bonthron, D.T.; Donlon, T.A.; Bruns, G.A.P.; Latt, S.A.
Science 228, 1401-1406, 1985
A:Title: Human von Willebrand factor (vWF): isolation of complementary DNA (cDNA) cto
A:Reference number: A44178; MUID:85244588
A:Accession: A44178
A:Molecule type: mRNA
A:Residues: 2621-2813 <GIN>
A:Cross-references: EMBL:X03028; NID:9340308; PIDN:AAA61293.1; PID:9340309
R:Verweij, C.L.; de Vries, C.J.M.; Distel, B.; van Zonneveld, A.J.; van Kessel, A.G.;
Nucleic Acids Res. 13, 4699-4717, 1985
A:Title: Construction of cDNA coding for human von Willebrand factor using antibody p
A:Reference number: S07363; MUID:85269603
A:Accession: S07363
A:Molecule type: mRNA
A:Residues: 2731-2813 <VE3>
A:Cross-references: EMBL:X02672; NID:937939; PIDN:CAA26503.1; PID:937940
R:Lynch, D.C.; Zimmerman, T.S.; Collins, C.J.; Brown, M.; Morin, M.J.; Ling, E.H.; Li
Cell 41, 49-56, 1985
A:Title: Molecular cloning of cDNA for human von Willebrand factor: authentication by
A:Reference number: S23678; MUID:85201687
A:Accession: S23678
A:Molecule type: mRNA
A:Residues: 2731-2813 <LYN>
A:Cross-references: EMBL:X03028
R:Titani, K.; Kumar, S.; Takio, K.; Ericsson, L.H.; Wade, R.D.; Ashida, K.; Walsh, K.
Biochemistry 25, 3171-3184, 1986
A:Title: Amino acid sequences of human von Willebrand factor.
A:Reference number: A90505; MUID:86269895
A:Accession: A90505
A:Molecule type: protein
A:Residues: 764-788, 'A', 790-1471, 'D', 1473-2813 <TT>
A:Note: 789-Thr was also found
R:Chopek, M.W.; Girma, J.P.; Fujikawa, K.; Davie, E.W.; Titani, K.
Biochemistry 25, 3146-3155, 1986
A:Title: Human von Willebrand factor: a multivalent protein composed of identical sub
A:Reference number: A23464; MUID:86269892
A:Accession: A23464
A:Molecule type: protein
A:Residues: 764-773, 2803-2813 <CHO>
R:Dent, J.A.; Berkowitz, S.D.; Ware, J.; Kaasper, C.K.; Ruggieri, Z.M.
Proc. Natl. Acad. Sci. U.S.A. 87, 6306-6310, 1990
A:Title: Identification of a cleavage site directing the immunochemical detection of
A:Reference number: A36013; MUID:90349604
A:Accession: A36013
A:Molecule type: protein
A:Residues: 1606-1617 <DEN>
R:Pay, P.J.; Kawel, Y.; Wagner, D.D.; Ginsburg, D.; Bonthron, D.; Ohlsson-Wilhelm, B.
Science 233, 995-998, 1986
A:Title: Propolypeptide of von Willebrand factor circulates in blood and is identical
A:Reference number: A60913; MUID:86208144
A:Accession: A60913
A:Molecule type: protein
A:Residues: 576-590 <FAV>
C:Genetics:
A:Gene: GDB:VWF
A:Cross-references: GDB:119125; OMIM:193400
A:Map position: 12p13.3-12p13.2
A:Initons: 19/1; 74/1; 108/2; 178/1; 219/3; 292/1; 333/1; 370/2; 386/1; 431/3; 478/1;
5/1; 1124/1; 1771/1; 1819/1; 1874/1; 1888/3; 1948/1; 2021/3; 2086/1; 2200/1; 2266/3;
C:Superfamily: von Willebrand factor; von Willebrand factor type A repeat homology; v
C:Keywords: blood coagulation; cell binding; connective tissue; disulfide bond; dupli
F:1-22/Domain: signal sequence *status predicted <SIG>
F:23-763/Product: von Willebrand antigen II *status predicted <MA>
F:34-386/Domain: type D repeat 1 <DD1>
F:387-745/Domain: type D repeat 2 <DD2>
F:698-700/Region: cell attachment (R-G-D) motif
F:764-2813/Product: von Willebrand factor *status predicted <MA2>
F:784-865/Domain: D *DDD>
F:788-833, 2216-2261/Region: duplication

F:826-853,2400-2515,2544-2662/Region: duplication
 F:842-1130,1934-2203/Region: duplication
 F:856-1241/Domain: type D repeat 3 <DD3>
 F:1375-1443/Domain: von Willebrand factor type A repeat homology <VWA1>
 F:1496-1654/Domain: von Willebrand factor type A repeat homology <VWA2>
 F:1689-1854/Domain: von Willebrand factor type A repeat homology <VWA3>
 F:1947-2295/Domain: type D repeat 4 <VD4>
 F:2296-2330/Domain: type B repeat 1 <VB1>
 F:2340-2365/Domain: type B repeat 2 <VB2>
 F:2375-2399/Domain: type B repeat 3 <VB3>
 F:2430-2497/Domain: von Willebrand factor type C repeat homology <VWC1>
 F:2507-2509/Domain: cell attachment (R-G-D) motif
 F:2581-2647/Domain: von Willebrand factor type C repeat homology <VWC2>
 F:857-1231,1515,1574,2223,2290,2357,2400,2546,2585,2790/Binding site: carbohydrate (Asn)
 F:1147/Binding site: carbohydrate (Asn) (covalent) #status atypical
 F:1248,1255,1256,1468,1477,1487,1679,2296/Binding site: carbohydrate (Thr) (covalent) #
 F:1263,1486/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 6.7% Score 110.5; DB 1; Length 2813;
 Best Local Similarity 21.0%; Pred. No. 3.3; Indels 91; Gaps 13;
 Matches 56; Conservative 30; Mismatches 90;

Qy 77 TPACDDPVTHMISFGSPSSVSSSTSSDRLDHPAPVADQASPRVRIOPKV 136
 Db 727 TMCCEDEDFM--CTMCGVPSLLPDVLSPP--LSH-----RSKRLSLCRPMV 772
 Qy 137 H-----CQPLDK-----GPAVPELDHFLCLCEACGCKC-- 167
 Db 773 KLVCPADNLRAEGLECKTKTCQNYDECMSCGVSGLCPGMVHNENRCVLALEKCPCHQ 832
 Qy 168 -KECASPTPLPSWCWNOECSCSQTLYVNGTCMLVGIYHCTNEDEGSCADHPC-- 224
 Db 833 GKRAVAGETV-----IGCNTVC-----NDRKKNCTDHYCDA 865
 Qy 225 SCSSNSCCARMSFGALSVLPCLLYL-----PATCGVLAORGYRLRRPGCRC 275
 Db 866 TCSITGMAYLYLFDG-LKYLPEGCOYLYVDYCGSNPGTRILYVGNK--CSHPVSK 921
 Qy 276 KHTNSVICKA-----SQDATSTRPK 297
 Db 922 KKRVTILVEGGEILELDEGVNKRPMK 948

RESULT 6

hypothetical protein DKFZP434H204.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T14764
 R:Manbut, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, August 1999
 A:Reference number: Z18181
 A:Accession: T14764
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-898 <MA8>
 A:Cross-references: EMBL:AL110226
 A:Experimental source: adult testis; clone DKFZP434H204
 C:Genetics:
 A:Note: DKFZP434H204.1

Query Match 6.6% Score 109; DB 2; Length 898;
 Best Local Similarity 23.5%; Pred. No. 1.4; Indels 118; Gaps 22;
 Matches 80; Conservative 18; Mismatches 125;

Qy 5 IPQSAPLTPNSVYQPLDLSRMSHRLQPLTILPIDQVKTSHVENDYID--NPSIALT 61
 Db 600 VPEIQPLAPSLAAGRPAD-----PLVYR-----NMSWQGNNSSECTTGLAV 644
 Qy 62 TGPKRTRGG-----APELAPPA-RCDDQVTHMISFGSPSSVSSSTSSDRLDHPM 115

Db 645 WRFVRCSSGDEDCAPARPPARRCHLRPCATW--HSGNWSCKNSCGGSSVVDVQ-- 700
 Qy 116 APPPVADQASPRVRIOPKVHQQPLDKGPVPE-----LDKHFLLC-EAC 162
 Db 701 ----CVDTDLRLRLR-----PFHCQP--GPAKPPAHRCCGAPCLSKWTSSHRECSEAC 748
 Qy 163 G-----KCK-----CKDCASPTPLPSC-----WVCNOCLCSAOTLYVNGTCMLV 203
 Db 749 GGGEGQRLVTCPEPGCEALRPNTTRPCNTHRCQTQWVGPWGQCSA-----PCGGGV 801
 Qy 204 QGIFNYCTN-----EDDEGSC-----ADHRC-----CSRNC-----CARW 235
 Db 802 QRLVLCVNTQTGLPEBDSQCEHAMPSSRCGTEDEPEVPPRCERDRLSFGCEPL 861
 Qy 236 SFMGALSVLP-----CLLCYLPATGCVKLAORGYRLR 270
 Db 862 RLIGRCQ--LPTRTQCNSCSPSHGA--PSRGHQRVAR 897

RESULT 7

crumbs protein - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 21-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 11-Jan-2000
 C:Accession: A35672
 R:Teppas, U.; Theres, C.; Knaul, E.
 Cell 61, 787-799, 1990
 A:File: crumbs encodes an EGF-like protein expressed on apical membranes of Drosophi
 A:Reference number: A35672; M0ID:90263104
 A:Accession: A35672
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-2139 <TEP>
 A:Cross-references: GB:M33753
 C:Genetics:
 A:Gene: FlyBase:crb
 A:Superfamily: unassigned EGF-related proteins; EGF homology
 C:KeyWords: transmembrane protein
 F:352-385/Domain: EGF homology <EGF1>
 F:392-424/Domain: EGF homology <EGF1>
 F:691-722/Domain: EGF homology <EGF1>
 F:767-799/Domain: EGF homology <EGF3>
 F:1878-1914/Domain: EGF homology <EGX2>

Query Match 6.5% Score 108; DB 2; Length 2139;
 Best Local Similarity 25.8%; Pred. No. 3.8; Indels 84; Gaps 15;
 Matches 63; Conservative 15; Mismatches 82;

Qy 68 RGAPELAPPA-RCDDQ-----VTHMISFGSPSSVSSSTSSDQ----- 109
 Db 1658 KGWESWDPPTPAFSTDIQDAAFQSLINTSYQVLYGMPESRQARGSTLSAQSSQFQGV 1717
 Qy 110 ---RLDHPAPP-PVADQASPRVRIOPKVHQCPLDKGPVPELDKFLFLLCEAGCK 165
 Db 1718 GEARLGLLLPYFSMAELYSRTNYSVOOKA-----QFRINATRPBEG-----CILCFOS 1766
 Qy 166 KCKE---CASPTPLPSCMV-----NOCLCSAOTLYVNGTCMLVYG 205
 Db 1767 DCKNDGFCQSPSDEYAC-TCQPGFSGDDCGTDDECLNTRCL-----NNGTCINQVAA 1818
 Qy 206 IF-----YHCTNDEDEGSCADHPCSGSRNC-----CARWSFGALSVLP 246
 Db 1819 FFCQCGQCFEGEQRCEQNIDE--CADQPCN-NGGNCITLIASYVDCPE-DYMGPGCVLK 1874
 Qy 247 CLLC 250
 Db 1875 QMTC 1878

RESULT 8

A, introns: 23/3; 71/3; 121/3; 183/1; 1083/3; 1141/3; 1408/1; 1449/1

Query Match 6.4%; Score 106.5; DB 2; Length 1513;
Best Local Similarity 21.0%; Pred. No. 3.6;
Matches 73; Conservative 38; Mismatches 123; Indels 113; Gaps 20;

OY 2 EPIPSAFLTPRSVWVQPLDLSRMSHRLQHPILPIDQVTSHEVDYIDNPDLALT 61
DB 1181 QPTVQGPAP---OLVQDP---YQTVQAPLSTVPVPAVTS-----AFSCA-- 1220
OY 62 TGRKRTRGCA-PELAPTPARCDDVT-----HHWISFGSPSSVSSSTSSDQL 111
DB 1221 SGQJITINFAVPECIFV---CEBSCTMCCEKPRQHCSCVSCQGTACATPAVQ-- 1275
OY 112 LDHMAPP-----VADQSPRAVNIQKRV-----HCQPL--DLK 144
DB 1276 ---AAPAPSCOPCOPAFCEPVCIAQOAP---VRIQINLATASSVLQASDACPFCBQSCV 1330
OY 145 GRAVPELDHFLLEACGCKCKEACSP-----RTLP-----SCWVCNOEC 186
DB 1331 QECQSTTLNVAATCOPACOSICQSCAPLGTSAFVWQITIPVVPVATAPVASTQLCAPKC 1390
OY 187 LCSAQTLYNVTGMCILVQGIIFYHCTNEDDEGSCADHPCSC---SRNCCARMSFGALS 242
DB 1391 ISDQGLCKNSPQC-IQGDASC-----QQLGTAPTPAVPLTVNVC----- 1433
OY 243 VLPC-----LCLYLPATGCVKLAQKCYDRLRRPGCKKTNSTVCK 284
DB 1434 -NLPCDOCTQOCYHQAFTCAPACASACE-AQCPVVSCEBACOTVCK 1478

RESULT 12

T42215
zonadhesin - mouse
N:Alternate names: sperm-specific membrane protein
C:Species: Mus musculus (house mouse)
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T42215
R:Gao, Z.; Garbers, D.L.
J. Biol. Chem. 273, 3415-3421, 1998
A:Title: Species diversity in the structure of zonadhesin, a sperm-specific membrane protein
A:Reference number: 222080; MUID:98123114
A:Accession: T42215
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-5376 <GAP>
A:Cross-references: EMBL:U97068; NID:q3327420; PID:q3327421; PIDN:AAC26680.1
C:Genetics:
A:Gene: Zan
A:Map position: 5
C:Function:
A:Description: functions in multiple cell adhesion processes
A>Note: found exclusively on the apical region of the sperm head
C:Keywords: cell adhesion

Query Match 6.3%; Score 105; DB 2; Length 5376;
Best Local Similarity 22.3%; Pred. No. 15;
Matches 65; Conservative 23; Mismatches 108; Indels 96; Gaps 15;

OY 54 DNPSIALTTPGKRRGCAPELAPPARCDQVTHHWSFGSPSSVSSSTSSDQLLD 113
DB 4079 DKCVLRIGCGKDGGLIAGRT-----WIS-----SDCTKSCSCMGIIICRD 4123
OY 114 HMAPPPVADQSPRAVNIQKRVKQPLDLKRAVPELDHFLLC-EAC-----GK 164
DB 4124 FQCPGTYCKESNDSSR-----CAKIPDQPA-----HSHYTNCLPACRSCTDLDDG 4172
OY 165 CKKCKASPRTPSCGVCNOECISAOITLVNVTGMCILVQGIIFYHCTNEDDEG----- 217
DB 4173 C---EGTSPK-VPS--PKCEGLQDPGVVYVNHKCVLDI-----HCGKDAQGGVVPAGK 4221

OY 218 -----SCADHPCSCSRNCCARMSFGALSVPCL 248
DB 4222 TWISRGCTGSCAGVAGVQCHNFCTPTGTCOONSCKSTVQCPAHSGYTTCLPCLPS- 4280
OY 249 LCLYLPATGCVKLAQKCYDRLRRPGCR-----KITNSVI-----CKAASD 289
DB 4281 -CFDEGLCGGASPRAPSTCRE-GCVCEADYVLRKDCVLTQCGCKDAQGD 4330

RESULT 13

A29072
pulmonary surfactant protein SP 18 precursor - dog (fragment)
C:Species: Canis lupus familiaris (dog)
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 20-Aug-1999
C:Accession: B29072; A29072
R:Hamgood, S.; Benson, B.J.; Schilling, J.; Damm, D.; Clements, J.A.; White, R.T.
Proc. Natl. Acad. Sci. U.S.A. 84, 66-70, 1987
A:Title: Nucleotide and amino acid sequences of pulmonary surfactant protein SP 18 an
A:Reference number: A29072; MUID:87092398
A:Accession: B29072
A:Molecule type: mRNA
A:Residues: 1-363 <HAMP>
A:Cross-references: GB:M15170; NID:g164077; PIDN:AAA30893.1; PID:g164078
A:Accession: A29072
A:Molecule type: protein
A:Residues: 182-210 <HA2>
C:Superfamily: pulmonary surfactant protein B; saposin repeat homology
F:1-14/Domain: signal sequence #status predicted <SIG>
F:15-180/Domain: propeptide #status predicted <PRP>
F:54-146/Domain: saposin repeat homology <SAP1>
F:180-267/Domain: saposin repeat homology <SAP2>
F:181-363/Product: pulmonary surfactant protein SP 18 #status experimental <MAT>
F:273-358/Domain: saposin repeat homology <SAP3>

Query Match 6.3%; Score 104; DB 2; Length 363;
Best Local Similarity 22.8%; Pred. No. 1.4;
Matches 54; Conservative 24; Mismatches 83; Indels 76; Gaps 11;

OY 130 RIQKVVHCOPLDLKGPAPV-----PELDKHFLLCEAGCKCK-----CKECA 171
DB 123 QINPKII-KCHLDLCKRGLPEPQSELSPLDLK-LIPELPGALQVGPHTDLSQEQ 180
OY 172 SPRTLPSCWYCNDE-----LCSAQTLYNVTGTCMC----- 201
DB 181 LPIPLPCWLCRTLRIRIQAMIPKGLAVTVGGVCHVPLVVGITQCLGERTVLLDA 240
OY 202 -----LVQGIIFYHCTNEDDEG-SCADHPCSCS-RSNCCARMSFGCA----- 240
DB 241 LIGRLPQLVCGILVLCSHDSAGPALASLPSEMSQESQCLCMFVTQAGNHSEQATP 300
OY 241 LSVVLPCLLCLYLPATGCVKLAQKCYDRLR--RPGCKKHTNSYTCAGASGDAKTSRP 295
DB 301 QAIRQACLSSMLDRKCECFVEQHMPRLQTLASGRDAHT---COALGACRTTFSR 354

RESULT 14

T27283
hypothetical protein Y64G10A.f - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27283
R:Ainscough, R.
submitted to the EMBL Data Library, September 1999
A:Reference number: 220336
A:Accession: T27283
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1620 <MII>
A:Cross-references: EMBL:AL110498; NID:e1542303; PIDN:CAB54471.1; CESP:Y64G10A.f
A:Experimental source: clone Y64G10A
C:Genetics:
A:Gene: CESP:Y64G10A.f

Query Match	6.3%;	Score 104;	DB 2;	Length 1620;
Best Local Similarity	26.5%;	Pred. No. 5.8;		
Matches 41; Conservative	18;	Mismatches 58;	Indels 38;	Gaps 10;

best local similarity 26.5%; Pred.No. 5.8;
Matches 41; Conservative 18; Mismatches 58; Indels 38; Gaps 10;

```
0Y      158  LCEA-CGRCKCKE-----CASPRITFSCWVGNQEC-----LCSAQTLVNYGTCM 200
          ::| | | | |         |   |   | | | |       : : |   | : |
Db      1371 MCDALSGKCFCKPBGHSGSDCKSGCVGRFGPD--CNQLCSENGGVCDST---GSCV 1423
```

QY 201 CLVGGIFHC---TNEDESSCADHPSCSRSNCCARWSPGALSVVLPCLLYLPATGC 257

Db 1424 CPEGYIGKCEIACQSDRGRGPTCEKICNENGGTCDR-----LTGQGRCLPGFTGMT-C 1476

```

258 VKLQRG-YDRLRRGCRCKHTNSVICKASGDAK 291
      :: | : : ||| : : | |:: |
Db 1477 NQVPEGRFGAGCKEKRCRANGH--CNASSGECK 1508

```

RESULT 15
B37057

C:Species: Cavia porcellus (guinea pig)
integlin beta-6 chain - guinea pig (fragment)

```
C:\Date: 15-Feb-1991 #sequence_revision
```

C;Accession: B37057

U. S. Biol. Chem. 265

A:Title: Complete amino acid sequence of a novel integrin beta subunit (beta

A;Reference number: A37057; MUID:90307659

A;Accession: B3/U5/
A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-577 <SHE>

A: Note: the authors translated the codon AAA for word 'all' as
A/Cross-references: GB:M35197; GB:J05522; NID:g191277; PIDN:AAA37043.1; PID:g553845

as PRO, ACG for residue 355 as Met, GAG for residue 363 as Thr, ACC for residue 364 as

glycyl-L-histidyl-L-histidine

Keywords: cell adhesion; cytoskeleton; transmembrane protein

Query match	6.2%;	Score	102.5;	DB	2;	Length	577;
Best Local Similarity	27.68;	Pred	No	3	8.		

Matches	45;	Conservative	12;	Mismatches

157 *****

137 ETEACGCKCKEASPRILPSCW---CNOECLCSAOTLVN-----YGTMC--LV 203

Db 409 LCGDNGCEGEC---VCRSGWTGEYCN--CTTSTDTCISDGTLCSGRGDVCVCGKCV 467

00' 364 CATTURCOWWCCCC

204 GGTTCCTNEDEG-----SCADHPCCSRSRNCARWSFMGALSVLPCLLCYLPATG 256

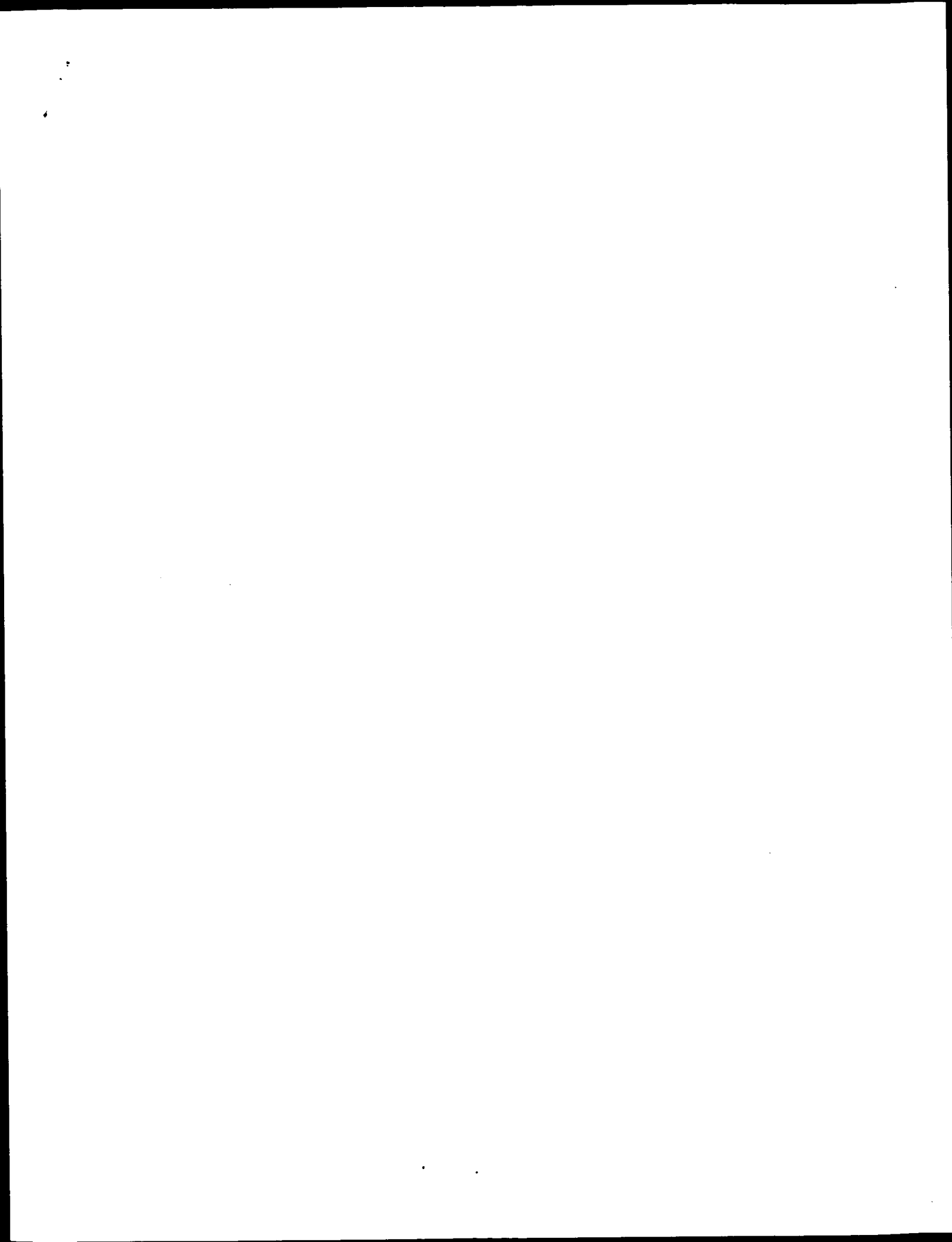
Db 463 -----CTNPGASGPTCERCPTCS-PCNSKRS-----CTECHI.SADG 499

2025 01 01 00:00:00

25 / CVKLAQRGYDLRRPG--C--RCKHTNSVICAA--SGDAKTS 293

Db 499 -----QPGEECVDKCKLAGVTISKADFSKDSVS 528

Search completed: May 8, 2002, 10:46:40
Job time: 52 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:45:48 ; Search time 8.33 Seconds
(without alignments)
640.295 million cell updates/sec

Title: US-10-082-902-2
Perfect score: 1657
Sequence: 1 MEPPIPQASAPLT PNSSVMQPSVICKASGDAKTSRPDKPF 299

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

```
Searched:      81999 seqs, 17838306 residues
Total number of hits satisfying chosen parameters:  81999
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database : Pending_Patents_AA_New:*
1 : /c9n2.6/plotdata/2/paa/PCr_NEW_COMB.pcp : *
2 : /c9n2.6/plotdata/2/paa/US06_NEW_COMB.pcp : *
3 : /c9n2.6/plotdata/2/paa/US07_NEW_COMB.pcp : *
4 : /c9n2.6/plotdata/2/paa/US08_NEW_COMB.pcp : *
5 : /c9n2.6/plotdata/2/paa/US09_NEW_COMB.pcp : *
6 : /c9n2.6/plotdata/2/paa/US10_NEW_COMB.pcp : *
7 : /c9n2.6/plotdata/2/paa/US10_NEW_COMB.pcp : *
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1657	100.0	299	1	PCT-US02-08781-8	Sequence 8, Appl11
2	1657	100.0	299	6	US-10-082-902-2	Sequence 2, Appl11
3	1621.5	97.9	375	1	PCT-US02-09921-1088	Sequence 2, Appl11
4	1554.5	93.8	300	6	US-10-082-902-5	Sequence 1088, Appl11
5	641.5	38.7	315	6	US-10-082-902-3	Sequence 5, Appl11
6	634.5	38.3	315	6	US-10-082-902-6	Sequence 3, Appl11
7	528	31.9	135	6	US-10-082-902-4	Sequence 6, Appl11
8	237	14.3	100	6	US-10-082-902-7	Sequence 4, Appl11
9	129	7.8	1432	6	US-10-123-155-255	Sequence 7, Appl11
10	121	7.3	2998	6	US-10-123-155-185	Sequence 255, Appl11
11	116	7.0	1162	6	US-10-123-155-313	Sequence 39, Appl11
12	114.5	6.9	3323	6	US-10-123-155-113	Sequence 185, Appl11
13	114	6.9	2477	6	US-10-123-155-331	Sequence 113, Appl11
14	112	6.8	2290	6	US-10-123-155-323	Sequence 331, Appl11
15	111.5	6.7	633	1	PCT-US02-09944-533	Sequence 323, Appl11
16	111.5	6.7	738	1	PCT-US02-09944-712	Sequence 533, Appl11
17	111.5	6.7	3552	6	US-10-123-155-339	Sequence 712, Appl11
18	111	6.7	2163	6	US-10-123-155-127	Sequence 339, Appl11
19	110.5	6.7	1871	6	US-10-063-545-91	Sequence 127, Appl11
20	110	6.6	2447	6	US-10-123-155-291	Sequence 91, Appl11
21	110	6.6	2542	6	US-10-123-155-991	Sequence 291, Appl11
22	109.5	6.6	771	6	US-10-123-155-495	Sequence 95, Appl11
23	109.5	6.6	1701	6	US-10-063-545-73	Sequence 495, Appl11
24	109	6.6	469	1	PCT-US02-09785-980	Sequence 73, Appl11
25	109	6.6	494	1	PCT-US02-09785-762	Sequence 980, Appl11
26	108.5	6.5	422	6	US-10-063-545-65	Sequence 762, Appl11
						Sequence 65, Appl11

45	106	6.4	1576	6	US-10-123-155-301	Sequence 301, App
44	106	6.4	1570	6	US-10-123-155-455	Sequence 455, App
43	106	6.4	1523	6	US-10-123-155-429	Sequence 429, App
42	106.5	6.4	2768	6	US-10-063-545-15	Sequence 15, App
41	106.5	6.4	1971	6	US-10-123-155-139	Sequence 139, App
40	107	6.5	3819	6	US-10-123-155-405	Sequence 405, App
39	107	6.5	2236	6	US-10-123-155-361	Sequence 361, App
38	107	6.5	1837	6	US-10-123-155-243	Sequence 243, App
37	107	6.5	1234	6	US-10-063-545-63	Sequence 63, App
36	107	6.5	804	6	US-10-123-155-149	Sequence 149, App
35	107.5	6.5	3907	6	US-10-029-217A-74	Sequence 24, App
34	107.5	6.5	3884	6	US-10-123-155-145	Sequence 145, App
33	107.5	6.5	2651	6	US-10-063-545-23	Sequence 23, App
32	107.5	6.5	1857	6	US-10-123-155-335	Sequence 365, App
31	108	6.5	2762	6	US-10-123-155-153	Sequence 13, App
30	108	6.5	1676	6	US-10-063-545-55	Sequence 53, App
29	108	6.5	1312	6	US-10-123-155-397	Sequence 397, App
28	108.5	6.5	3479	6	US-10-123-155-123	Sequence 123, App
27	108.5	6.5	2849	6	US-10-123-155-311	Sequence 371, App

ALIGNMENTS

```

RESULT 1
PCT-US02-08781-8
: Sequence 8, Application PC/TUS0208781
: GENERAL INFORMATION:
: APPLICANT: Hyseq, Inc
: APPLICANT: Tang, Tom Y.
: APPLICANT: Zhou, Ping
: APPLICANT: Goodrich, Ryle
: APPLICANT: Asundi, Vinod
: APPLICANT: Liu, Chenghua
: APPLICANT: Wehrman, Tom
: APPLICANT: Ren, Feiyan
: APPLICANT: Dirmanac, Radoje, T.
: TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
: FILE REFERENCE: 21272-006-061 CIP
: CURRENT APPLICATION NUMBER: PCT/US02/08781
: PRIOR FILING DATE: 2002-03-20
: PRIOR APPLICATION NUMBER: 09/814,354
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: Patentin version 3.0
: SEQ ID NO 8
: LENGTH: 299
: TYPE: PRT
: ORGANISM: Homo sapiens
PCT-US02-08781-8

Query Match 100.0%; Score 1657; DB 1; Length 299;
Best Local Similarity 100.0%; Pred. No. 4,2e-115;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 1 MEPPIQSAPLTPLNSVMVVOPLDLSHMSHSRLQHPLTLLIPIDVKTSHENDYIDNPSLAL 60
Db 1 MEPPIQSAPLTPLNSVMVVOPLDLSHMSHSRLQHPLTLLIPIDVKTSHENDYIDNPSLAL 60

QY 61 TTGPKRTKGAPLAPTPARCDQDVTNHWISFGSPSSVSSSSSTSSDQRLDHHKAPPV 120
Db 61 TTGPKRTKGAPLAPTPARCDQDVTNHWISFGSPSSVSSSSSTSSDQRLDHHKAPPV 120

QY 121 ADQASPRAVRITQPKVYHCOPLDLKGPAPRPPELDKHFLLCGAGKCKCEKASPRLLPSCW 180
Db 121 ADQASPRAVRITQPKVYHCOPLDLKGPAPRPPELDKHFLLCGAGKCKCEKASPRLLPSCW 180

QY 181 VCNQGLCSAQTIVLVYGCMLVQGIFFHCTNEBDEGSCADPCSCSINCCARMSFMCA 240
Db 181 VCNQGLCSAQTIVLVYGCMLVQGIFFHCTNEBDEGSCADPCSCSINCCARMSFMCA 240

QY 241 LSVYLPCLLCVLPATGCVKLAQGRYDRLRRPQCRCKHTNSYICAKASGDAKTSRDPKP 299

```

Db 241 LSVVLPCLCYLPATGCVKLAORGYDLRRPGCRCKHTNSVYCKAASGDATSRDPKPF 299

RESULT 2

US-10-082-902-2
Sequence 2, Application US/10082902
GENERAL INFORMATION:
APPLICANT: Ballinger, Dennis G.
APPLICANT: Montgomery, Julie R.
TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
FILE REFERENCE: 28110/35878
CURRENT APPLICATION NUMBER: US/10/082, 902
CURRENT FILING DATE: 2002-02-26
PRIOR APPLICATION NUMBER: US/09/370, 398
PRIOR FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 299
TYPE: PRT
ORGANISM: Homo sapiens
US-10-082-902-2

Query Match 100.0%; Score 1657; DB 6; Length 299;
Best Local Similarity 100.0%; Pred. No. 4.2e-115;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEPPPOSAPLTPNSVWVQPLDLSRMSHSLQHPILTLPIQVYTSVHENDYIDNPSLAL 60
Db 1 MEPPPOSAPLTPNSVWVQPLDLSRMSHSLQHPILTLPIQVYTSVHENDYIDNPSLAL 60
Qy 61 TTGPKRTGGAPELAPTPARCDDVTHHMFSGRPSVSSSSSTSDQRLDHPAPPV 120
Db 61 TTGPKRTGGAPELAPTPARCDDVTHHMFSGRPSVSSSSSTSDQRLDHPAPPV 120
Qy 121 ADQASPRAVRIQPKVHCOPLDLKGPVAPPELDKHFLLCEAGCKCKECCASPRTLPSCW 180
Db 121 ADQASPRAVRIQPKVHCOPLDLKGPVAPPELDKHFLLCEAGCKCKECCASPRTLPSCW 180
Qy 181 VCNQECLSAQTLYNVTGCMCLVQGIFFHCTNEDEGSCADHPSCSSNCCARMSFMGA 240
Db 181 VCNQECLSAQTLYNVTGCMCLVQGIFFHCTNEDEGSCADHPSCSSNCCARMSFMGA 240
Qy 241 LSVVLPCLCYLPATGCVKLAORGYDLRRPGCRCKHTNSVYCKAASGDATSRDPKPF 299
Db 241 LSVVLPCLCYLPATGCVKLAORGYDLRRPGCRCKHTNSVYCKAASGDATSRDPKPF 299

RESULT 3

PCT-US02-09921-1088
Sequence 1088, Application PC/TUS0209921
GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: DAFEO, Abel
APPLICANT: JONES, Anissa L.
APPLICANT: TRAN, Alanna-Phung B.
APPLICANT: DAHL, Christopher R.
APPLICANT: GIETZEN, Darryl
APPLICANT: CHINN, Joyce
APPLICANT: DUFOUR, Gerard E.
APPLICANT: HILLMAN, Jennifer L.
APPLICANT: YU, Jimmy Y.
APPLICANT: TUASON, Olivia
APPLICANT: VAP, Pierre E.
APPLICANT: AMSHER, Stefan R.
APPLICANT: DAUGHERTY, Sean C.
APPLICANT: DAM, Tammy C.
APPLICANT: LIU, Tommy F.
APPLICANT: NGUYEN, Duy-Viet An
APPLICANT: KLEPFELD, Joel
APPLICANT: GERSTIN JR., Edward H.

APPLICANT: PERALTA, Careyna H.
APPLICANT: DAVID, Marie H.
APPLICANT: LEWIS, Samantha A.
APPLICANT: CHEN, Alice J.
APPLICANT: PANZER, Scott R.
APPLICANT: HARRIS, Bernard
APPLICANT: FLORES, Vincent
APPLICANT: MARMAHA, Rakesh
APPLICANT: LO, Audrey
APPLICANT: IAN, Ruth Y.
APPLICANT: URASHKA, Michael
TITLE OF INVENTION: SECRETORY MOLECULES
FILE REFERENCE: PT-1232 PCT
CURRENT APPLICATION NUMBER: PCT/US02/09921
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: 60/280, 067; 60/280, 068; 60/291, 280; 60/291, 849;
60/291, 829; 60/299, 428; 60/300, 001; 60/299, 776
PRIOR FILING DATE: 2001-03-29; 2001-03-29; 2001-05-16; 2001-05-17;
2001-05-17; 2001-06-19; 2001-06-20; 2001-06-20
NUMBER OF SEQ ID NOS: 1146
SOFTWARE: PERL Program
SEQ ID NO 1088
LENGTH: 375
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: LG:241742.1.orf3:2001JUN22
PCT-US02-09921-1088

Query Match 97.9%; Score 1621.5; DB 1; Length 375;
Best Local Similarity 97.7%; Pred. No. 2.2e-112;
Matches 293; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

Qy 1 MEPPPOSAPLTPNSVWVQPLDLSRMSHSLQHPILTLPIQVYTSVHENDYIDNPSLAL 60
Db 76 MEPPPOSAPLTPNSVWVQPLDLSRMSHSLQHPILTLPIQVYTSVHENDYIDNPSLAL 135
Qy 61 TTGPKRTGGAPELAPTPARCDDVTHHMFSGRPSVSSSSSTSDQRLDHPAPPV 119
Db 61 TTGPKRTGGAPELAPTPARCDDVTHHMFSGRPSVSSSSSTSDQRLDHPAPPV 195
Qy 136 TTGPKRTGGAPELAPTPARCDDVTHHMFSGRPSVSSSSSTSDQRLDHPAPPV 195
Db 136 TTGPKRTGGAPELAPTPARCDDVTHHMFSGRPSVSSSSSTSDQRLDHPAPPV 195
Qy 120 VADQASPRAVRIQPKVHCOPLDLKGPVAPPELDKHFLLCEAGCKCKECCASPRTLPSCW 179
Db 120 VADQASPRAVRIQPKVHCOPLDLKGPVAPPELDKHFLLCEAGCKCKECCASPRTLPSCW 255
Qy 180 VCNQECLSAQTLYNVTGCMCLVQGIFFHCTNEDEGSCADHPSCSSNCCARMSFMGA 239
Db 180 VCNQECLSAQTLYNVTGCMCLVQGIFFHCTNEDEGSCADHPSCSSNCCARMSFMGA 315
Qy 240 LSVVLPCLCYLPATGCVKLAORGYDLRRPGCRCKHTNSVYCKAASGDATSRDPKPF 299
Db 316 LSVVLPCLCYLPATGCVKLAORGYDLRRPGCRCKHTNSVYCKAASGDATSRDPKPF 375

RESULT 4

US-10-082-902-5
Sequence 5, Application US/10082902
GENERAL INFORMATION:

APPLICANT: Ballinger, Dennis G.
APPLICANT: Montgomery, Julie R.
TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
FILE REFERENCE: 28110/35878
CURRENT APPLICATION NUMBER: US/10/082, 902
CURRENT FILING DATE: 2002-02-26
PRIOR APPLICATION NUMBER: US/09/370, 398
PRIOR FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 300
TYPE: PRT

ORGANISM: Mus musculus
US-10-082-902-5

Query Match	93.88;	Score 1554.5;	DB 6;	Length 300;
Best Local Similarity	92.78;	Pred. NO. 1.4e-107;		
Matches 278;	Conservative 9;	Mismatches 12;	Indels 1;	Gaps 1.

[illegible]

RESULT 5
US-10-082-902-3
; Sequence 3, Application US/10082902
; General Information

```

1  APPLICANT: Ballinger, Dennis G.
2  APPLICANT: Montgomery, Julie R.
3  TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
4  FILE REFERENCE: 28110/35878
5  CURRENT APPLICATION NUMBER: US/10/082,902
6  CURRENT FILING DATE: 2002-02-26
7  PRIOR APPLICATION NUMBER: US/09/370,398
8  PRIOR FILING DATE: 1998-08-06
9  NUMBER OF SEQ ID NOS: 13
10 SOFTWARE: PatentIn Ver. 2.0
11 SEQ ID NO 3
12     LENGTH: 315
13     TYPE: PRT
14 ORGANISM: Homo sapiens
15 -S-10--082-902-3

```

Query Match	38.7%;	Score 641.5;	DB 6;	length 315;
Best Local Similarity	43.6%;	Pred. No. 1.9e-40;		
Matches 130; Conservative	45;	Mismatches 84;	Indels 39;	Gaps 8

[illegible]

RESULT 6
US-10-082-902-6
; Sequence 6, Application US/10082902
CENTRAL INTELLIGENCE

```

1  APPLICANT: Ballinger, Dennis G.
2  APPLICANT: Montgomery, Julie R.
3  TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
4  FILE REFERENCE: 28110/35878
5  CURRENT APPLICATION NUMBER: US/10/082,902
6  CURRENT FILING DATE: 2002-02-26
7  PRIOR APPLICATION NUMBER: US/09/370,398
8  PRIOR FILING DATE: 1998-08-06
9  NUMBER OF SEQ ID NOS: 13
10 SOFTWARE: Patentin Ver. 2.0
11 SEQ ID NO 6
12     LENGTH: 315
13     TYPE: PRT
14 ORGANISM: Mus musculus
15 US-10-082-902-6

```

Query Match	38.38;	Score 634.5;	DB 6;	Length 315;
Best Local Similarity	43.38;	Pred. No. 6.1e-40;		
Matches 129; Conservative	46;	Mismatches 84;	Indels 39;	Gaps 9

```

OY 19 QPILDSRMHSR-----IQPHLLIPIDOVKSHYENDYIDNPSLATLTGK--- 65
    ||||| : : : : :
Db 12 QPILQAHDSGROGEBPRDPAITDOOVHVLSDIDIRIRMTNETGCPVPRPGKRAP 71
    ||||| : : : : :
OY 66 -----RTG-----GAPELAPT--PARCDDYTHMISFGSRPSSVSSSTSD 10
    : : : : :
Db 72 RPSTQKHRLHGLPEHROPRLPDSOVHSAFSLRSISTVSSGSRSSYRPTSTSSSE 13
    : : : : :
OY 109 QRL--DHAAPRVDAQSPRAVRIQRVVHCQPLDLKGAAPVELDKPFLCEAGCKK 16
    ||||| : : : : :
Db 132 QRLGSGTSHGCAAD-----GIIRQRP--SELKAGDIK-PLSKDDLGILAHYRCEGCKK 18
    ||||| : : : : :
OY 167 CKECASPTLPSCWVNOBCELSAQTLVNTGTCMLQGIIFYHCTNEDDEGSCADHPSC 22
    ||||| : : : : :
Db 186 CKECTYPRPLPSDMIDCKOCGCSAQNVNIDYGTGVCCVKGFLYHCNS-DEEDNCADNPSC 24
    ||||| : : : : :
OY 227 SRNCCARSPFGALSVYLPCLLCTYLPAATGCVKALQAGYDRLRPPCCRCCKHTNSYICK 264
    ||||| : : : : :
Db 245 SOSHCCTTNSAGVNSLPLPCLMCKYLAAGKGLCKQCCQDVRNRPCCRCCKNSNTYOCK 302
    ||||| : : : : :

```

RESULT 7
US-10-082-902-4
; Sequence 4, Application US/10082902
: GENERAL INFORMATION.

```

? APPLICANT: Montgomery, Julie R.
? TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
? FILE REFERENCE: 28110/35878
? CURRENT APPLICATION NUMBER: US/10/082,902
? CURRENT FILING DATE: 2002-02-26
? PRIOR APPLICATION NUMBER: US/09/370,398
? PRIOR FILING DATE: 1998-08-06
? NUMBER OF SEQ ID NOS: 13
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 4
? LENGTH: 139
? TYPE: prt
? ORGANISM: Homo sapiens
? OS-10--082-902-4

```

Query Match	31.98;	Score 528;	DB 6;	Length 139;
Best Local Similarity	-62.68;	Pred. NO. 1.7e-32;		
Matches	87;	Conservative	24;	Mismatches 24;
			Indels	4;
			Gaps	1

```
OY 159 CEAGCKCKCKEASPRTPSCWVNOECICSAOTLVNVTGCMCLVGIHFYHCTNEDEGS 218
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 255
; LENGTH: 1432
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-255

Query Match
Best Local Similarity 26.6%; Pred. No. 0.044; Indels 24; Gaps 6;
Matches 38; Conservative 10; Mismatches 71;

OY 159 CEAGCKCKCKEASPRTPSCWVNOECICSAOTLVNVTGCMCLVGIHFYHCTNEDEGS 218
DB 583 ctacagctctccaataatgacacagcttggagtggttgga--acaggaact 639
OY 219 CA-----DHPGCSRSRNCARMSFGALSVVLPCLLCYLPATGCVKLAORGYDRLR 270
DB 640 caaccagcaggtctctatccaatggtgcctgt-----ccaggtacagcatgt----- 688
OY 271 PGCRKHTNSVICKAASGDAKTS 293
DB 689 -ac-caattgcacatacgaaatt 709

RESULT 10
US-10-123-155-39
; Sequence 39, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 39
; LENGTH: 2998
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-39

Query Match
Best Local Similarity 7.3%; Score 121; DB 6; Length 2998;
Matches 40; Conservative 7; Mismatches 60; Indels 26; Gaps 7;

OY 161 ACGRCKCKEASPR---TLPSCWVNOECICSAOTLVNVTGCMCLVGIHFYHCTNEDEGS 217
DB 562 acgaactctgataatctctgctctatccaagt---gacacgatagc--ct----- 610
OY 218 SCADHPGCSRSRNCARMSFGALSVVLPCLLCYLPATGCVKLAORGYDRLRPGCRCKH 277
DB 611 tca---cact---ccactggagtcgccgctctccctctgcataga-----agttga 658

OY 108 DORLDHNAAPPVADQASPRRAVRIOPKVVHCQPLDLKGPAVPELD----- 153
DB 5 DORLASTPSP-----SGQSIIRTP-----GAGVHPKADGALKGEAEGSAGHP 49
OY 154 -KHPLCEAGCKCKCKEASPRTPSCWVNOECICSAOTLVNVTGCMCLV 203
DB 50 SEHLFICEGCKCKCVPTAARPLPSWMLCQRCICSAESLDDYGTCLCV 100

Query Match
Best Local Similarity 14.3%; Score 237; DB 6; Length 100;
Matches 44; Conservative 14; Mismatches 23; Indels 30; Gaps 3;

OY 108 DORLDHNAAPPVADQASPRRAVRIOPKVVHCQPLDLKGPAVPELD----- 153
DB 5 DORLASTPSP-----SGQSIIRTP-----GAGVHPKADGALKGEAEGSAGHP 49
OY 154 -KHPLCEAGCKCKCKEASPRTPSCWVNOECICSAOTLVNVTGCMCLV 203
DB 50 SEHLFICEGCKCKCVPTAARPLPSWMLCQRCICSAESLDDYGTCLCV 100

RESULT 9
US-10-123-155-255
; Sequence 255, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT FILING DATE: 2002-04-15
```

```
OY 159 CEAGCKCKCKEASPRTPSCWVNOECICSAOTLVNVTGCMCLVGIHFYHCTNEDEGS 218
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 255
; LENGTH: 1432
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-255

Query Match
Best Local Similarity 7.8%; Score 129; DB 6; Length 1432;
Matches 38; Conservative 10; Mismatches 71; Indels 24; Gaps 6;

OY 159 CEAGCKCKCKEASPRTPSCWVNOECICSAOTLVNVTGCMCLVGIHFYHCTNEDEGS 218
DB 583 ctacagctctccaataatgacacagcttggagtggttgga--acaggaact 639
OY 219 CA-----DHPGCSRSRNCARMSFGALSVVLPCLLCYLPATGCVKLAORGYDRLR 270
DB 640 caaccagcaggtctctatccaatggtgcctgt-----ccaggtacagcatgt----- 688
OY 271 PGCRKHTNSVICKAASGDAKTS 293
DB 689 -ac-caattgcacatacgaaatt 709

RESULT 10
US-10-123-155-39
; Sequence 39, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 39
; LENGTH: 2998
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-39

Query Match
Best Local Similarity 7.3%; Score 121; DB 6; Length 2998;
Matches 40; Conservative 7; Mismatches 60; Indels 26; Gaps 7;

OY 161 ACGRCKCKEASPR---TLPSCWVNOECICSAOTLVNVTGCMCLVGIHFYHCTNEDEGS 217
DB 562 acgaactctgataatctctgctctatccaagt---gacacgatagc--ct----- 610
OY 218 SCADHPGCSRSRNCARMSFGALSVVLPCLLCYLPATGCVKLAORGYDRLRPGCRCKH 277
DB 611 tca---cact---ccactggagtcgccgctctccctctgcataga-----agttga 658
```

QY 278 ENSVICKAASGA 290
 Db 659 ttctacaaagaa 671

RESULT 11

US-10-123-155-185

Sequence 185, Application US/10123155

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: Deforge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tamas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Stewart, Timothy A.

APPLICANT: Tamas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

Query Match 6.9%; Score 114.5; DB 6; Length 3323;

Best Local Similarity 27.2%; Pred. No. 1.3;

Matches 41; Conservative 7; Mismatches 62; Indels 41; Gaps 9;

QY 159 CEACGCKCKEASPRTPSCWVNCNOELCSAQTLYVNGTC-----M 200

Db 1597 caccgctc--ccgtgcatccacgacacacactgtggtcgacgctccagcgtcaaa 1654

QY 201 CLVGGIFVHCINDEGSCADHPCGSCSRNCCARMSFGALSVLPCLLCYLPATGCKYL 260

Db 1655 caaa-----cag-----gacaaactcagttccatg---ga-----acttcc--ttccga 1696

QY 261 AORGYDRLRPGCRCKHTNSVI-CKAASGA 290

Db 1697 atg--acttgcacacacacacgaatgaa 1725

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

US-10-123-155-185

RESULT 13

US-10-123-155-331

Sequence 331, Application US/10123155

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: Deforge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tamas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

Query Match 6.9%; Score 114; DB 6; Length 2477;

Best Local Similarity 31.3%; Pred. No. 1;

Matches 41; Conservative 7; Mismatches 62; Indels 41; Gaps 9;

QY 159 CEACGCKCKEASPRTPSCWVNCNOELCSAQTLYVNGTC-----M 200

Db 1597 caccgctc--ccgtgcatccacgacacacactgtggtcgacgctccagcgtcaaa 1654

QY 201 CLVGGIFVHCINDEGSCADHPCGSCSRNCCARMSFGALSVLPCLLCYLPATGCKYL 260

Db 1655 caaa-----cag-----gacaaactcagttccatg---ga-----acttcc--ttccga 1696

QY 261 AORGYDRLRPGCRCKHTNSVI-CKAASGA 290

Db 1697 atg--acttgcacacacacgaatgaa 1725

RESULT 13

US-10-123-155-331

Sequence 331, Application US/10123155

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: Deforge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tamas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

APPLICANT: Zhang, Zemin

</

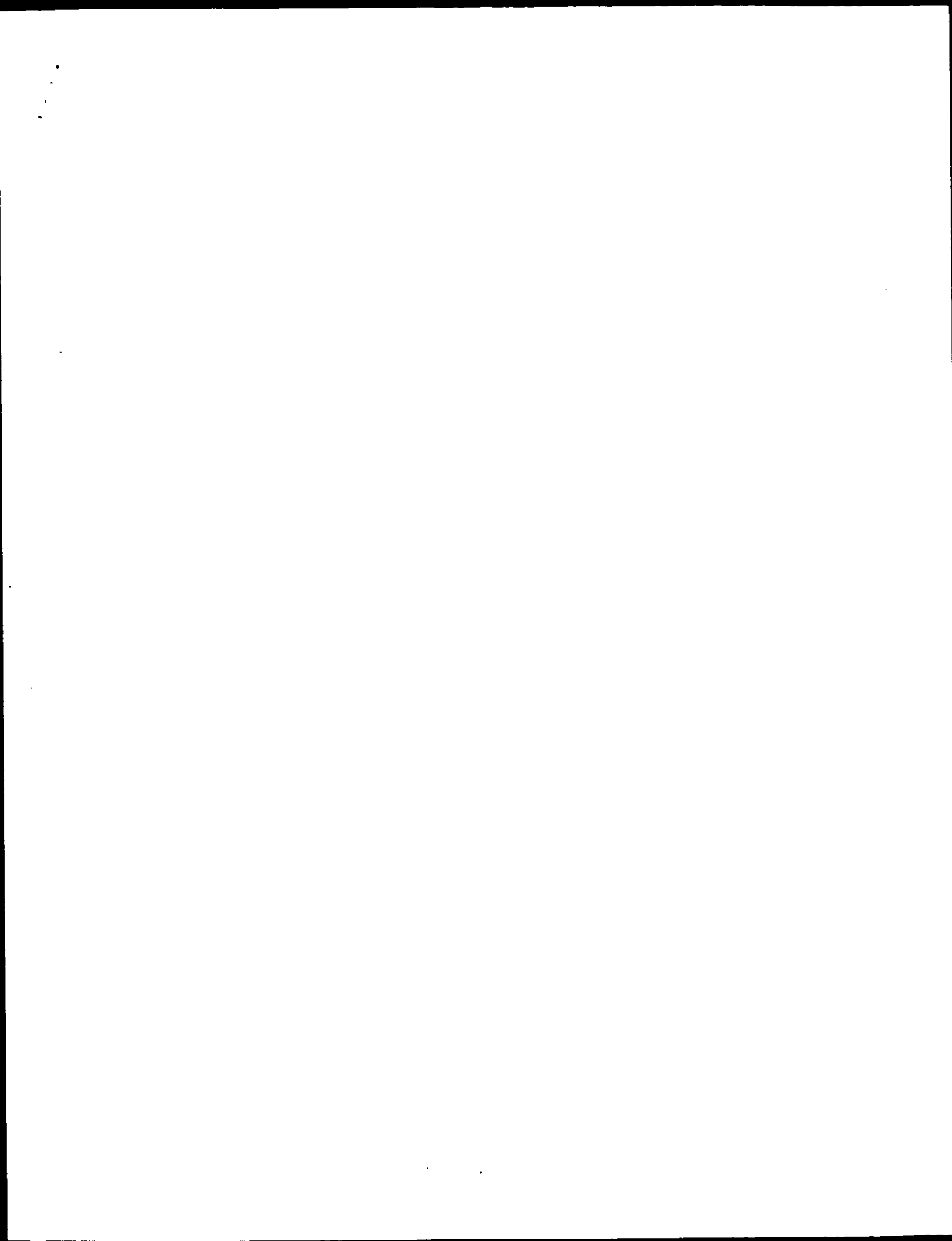
Sat May 11 12:51:25 2002

us-10-082-902-2.rapn

Page 7

Db 630 Ti 631

Search completed: May 8, 2002, 10:49:05
Job time: 197 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:45:48 ; Search time 105.96 Seconds

(without alignments)
991.176 Million cell updates/sec

Title: US-10-082-902-2

Sequence: 1 MEPPIPQAPLTPNSVWVQ.....SVICKAASGAKTSHPDKRF 299

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3516493 seqs, 351254056 residues

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/2/paa/PCRUS_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/2/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/2/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/2/paa/US103_COMB.pep.*
28: /cgn2_6/ptodata/2/paa/US104_COMB.pep.*
29: /cgn2_6/ptodata/2/paa/US105_COMB.pep.*
30: /cgn2_6/ptodata/2/paa/US106_COMB.pep.*
31: /cgn2_6/ptodata/2/paa/US107_COMB.pep.*
32: /cgn2_6/ptodata/2/paa/US108_COMB.pep.*
33: /cgn2_6/ptodata/2/paa/US109_COMB.pep.*
34: /cgn2_6/ptodata/2/paa/US110_COMB.pep.*
35: /cgn2_6/ptodata/2/paa/US111_COMB.pep.*
36: /cgn2_6/ptodata/2/paa/US112_COMB.pep.*
37: /cgn2_6/ptodata/2/paa/US113_COMB.pep.*
38: /cgn2_6/ptodata/2/paa/US114_COMB.pep.*
39: /cgn2_6/ptodata/2/paa/US115_COMB.pep.*
40: /cgn2_6/ptodata/2/paa/US116_COMB.pep.*
41: /cgn2_6/ptodata/2/paa/US117_COMB.pep.*
42: /cgn2_6/ptodata/2/paa/US118_COMB.pep.*
43: /cgn2_6/ptodata/2/paa/US119_COMB.pep.*
44: /cgn2_6/ptodata/2/paa/US120_COMB.pep.*
45: /cgn2_6/ptodata/2/paa/US121_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1657	100.0	299	17	US-09-370-398-2	Sequence 2, Appl1

2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45																																																																																																																																																																
1657	100.0	299	26	US-10-083-215-2	Sequence 2, Appl1	1657	100.0	299	26	US-10-090-190-2	Sequence 2, Appl1	1554.5	93.8	300	17	US-09-370-398-5	Sequence 5, Appl1	1554.5	93.8	300	26	US-10-083-215-5	Sequence 5, Appl1	1444.5	87.1	266	32	PCT-US01-27760A-819	Sequence 819, App	1444	87.1	266	32	PCT-US01-27760A-819	Sequence 819, App	679.5	41.0	319	18	US-09-488-725A-1833	Sequence 9, Appl1	679.5	41.0	319	24	US-60-139-556-9	Sequence 9, Appl1	679.5	41.0	319	26	US-10-018-170-9	Sequence 9, Appl1	679.5	41.0	319	18	US-09-488-725A-5405	Sequence 9, Appl1	679.5	41.0	319	17	US-09-393-314B-7	Sequence 7, Appl1	665.5	40.2	319	17	US-09-393-314B-24	Sequence 24, Appl1	654	39.5	294	17	PCT-US99-20638-3	Sequence 24, Appl1	653.5	39.4	288	17	PCT-US99-20638-2	Sequence 3, Appl1	653.5	39.4	288	17	US-09-393-314B-3	Sequence 3, Appl1	653.5	39.4	288	17	US-09-393-314B-2	Sequence 2, Appl1	653.5	39.4	288	17	US-09-393-314B-1	Sequence 1, Appl1	650	39.2	295	17	US-09-393-314B-5	Sequence 5, Appl1	650	39.2	295	17	US-09-393-314B-16	Sequence 16, Appl1	650	39.2	295	17	US-09-393-314B-15	Sequence 15, Appl1	650	39.2	295	17	US-09-393-314B-14	Sequence 14, Appl1	650	39.2	295	17	US-09-393-314B-13	Sequence 13, Appl1	650	39.2	295	17	US-09-393-314B-12	Sequence 12, Appl1	650	39.2	295	17	US-09-393-314B-11	Sequence 11, Appl1	650	39.2	295	17	US-09-393-314B-10	Sequence 10, Appl1	650	39.2	295	17	US-09-393-314B-9	Sequence 9, Appl1	650	39.2	295	17	US-09-393-314B-8	Sequence 8, Appl1	650	39.2	295	17	US-09-393-314B-7	Sequence 7, Appl1	650	39.2	295	17	US-09-393-314B-6	Sequence 6, Appl1	650	39.2	295	17	US-09-393-314B-5	Sequence 5, Appl1	650	39.2	295	17	US-09-393-314B-4	Sequence 4, Appl1	650	39.2	295	17	US-09-393-314B-3	Sequence 3, Appl1	650	39.2	295	17	US-09-393-314B-2	Sequence 2, Appl1	650	39.2	295	17	US-09-393-314B-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-370-398-2
Sequence 2, Application US/09370398
GENERAL INFORMATION:
APPLICANT: Ballinger, Dennis G.
TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
FILE REFERENCE: 28110/35878
CURRENT APPLICATION NUMBER: US/09/370, 398
CURRENT FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 299
TYPE: PRT
ORGANISM: Homo sapiens
US-09-370-398-2

Query Match 100.0%; Score 1657; DB 17; Length 299;
Best Local Similarity 100.0%; Pred. No. 3, 5e-125;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 MEPPIPQAPLTPNSVWVQPLDLSRMSHRLQHPITLPIQDKTSHVENDYIDNPSLAL 60

```

Db      1 MEPPIPQSAFLTPNSVWVQPLDLSRMSHSLQHPLTLPIQVKTSHVENDYIDNPSLAL 60
Qy      61 TTGPKRTGAGAPLAPTPACDODVTHHMTSFGSRPSSVSSSSSTSDQRLDHHMAPPV 120
Db      61 TTGPKRTGAGAPLAPTPACDODVTHHMTSFGSRPSSVSSSSSTSDQRLDHHMAPPV 120
Qy      121 ADQASPRAVRIQPKVHHCQPLDLKGPVAPPDLKHFLCEACGCKCKECASPTLPSCW 180
Db      121 ADQASPRAVRIQPKVHHCQPLDLKGPVAPPDLKHFLCEACGCKCKECASPTLPSCW 180
Qy      181 VCNQECLSAQTLVNYGTCMCLVQGIIFYHCTNEDDEGSCADHPCSCSRNSCCARMSFMGA 240
Db      181 VCNQECLSAQTLVNYGTCMCLVQGIIFYHCTNEDDEGSCADHPCSCSRNSCCARMSFMGA 240
Qy      241 LSVLPLCLCYLPATGCVKLAQRGYDRLRRPGCRCKHTNSVICKAASGDAKTSRPDKPF 299
Db      241 LSVLPLCLCYLPATGCVKLAQRGYDRLRRPGCRCKHTNSVICKAASGDAKTSRPDKPF 299

RESULT 2
US-10-083-215-2
; Sequence 2, Application us/10083215
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Montgomery, Julie R.
; TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
; FILE REFERENCE: 28110/35878
; CURRENT APPLICATION NUMBER: us/10/083, 215
; CURRENT FILING DATE: 2002-02-26
; PRIOR APPLICATION NUMBER: us/09/370, 398
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-083-215-2

```

```

Query Match      100.0%; Score 1657; DB 26; Length 299;
Best Local Similarity 100.0%; Pred. No. 3.5e-125;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MEPPIPQSAFLTPNSVWVQPLDLSRMSHSLQHPLTLPIQVKTSHVENDYIDNPSLAL 60
Db      1 MEPPIPQSAFLTPNSVWVQPLDLSRMSHSLQHPLTLPIQVKTSHVENDYIDNPSLAL 60
Qy      61 TTGPKRTGAGAPLAPTPACDODVTHHMTSFGSRPSSVSSSSSTSDQRLDHHMAPPV 120
Db      61 TTGPKRTGAGAPLAPTPACDODVTHHMTSFGSRPSSVSSSSSTSDQRLDHHMAPPV 120
Qy      121 ADQASPRAVRIQPKVHHCQPLDLKGPVAPPDLKHFLCEACGCKCKECASPTLPSCW 180
Db      121 ADQASPRAVRIQPKVHHCQPLDLKGPVAPPDLKHFLCEACGCKCKECASPTLPSCW 180
Qy      181 VCNQECLSAQTLVNYGTCMCLVQGIIFYHCTNEDDEGSCADHPCSCSRNSCCARMSFMGA 240
Db      181 VCNQECLSAQTLVNYGTCMCLVQGIIFYHCTNEDDEGSCADHPCSCSRNSCCARMSFMGA 240
Qy      241 LSVLPLCLCYLPATGCVKLAQRGYDRLRRPGCRCKHTNSVICKAASGDAKTSRPDKPF 299
Db      241 LSVLPLCLCYLPATGCVKLAQRGYDRLRRPGCRCKHTNSVICKAASGDAKTSRPDKPF 299

RESULT 3
US-10-090-190-2
; Sequence 2, Application us/10090190
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Montgomery, Julie R.
; TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods

```

```

; FILE REFERENCE: 28110/35878
; CURRENT APPLICATION NUMBER: us/10/090, 190
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: us/09/370, 398
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-090-190-2

```

```

Query Match      100.0%; Score 1657; DB 26; Length 299;
Best Local Similarity 100.0%; Pred. No. 3.5e-125;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MEPPIPQSAFLTPNSVWVQPLDLSRMSHSLQHPLTLPIQVKTSHVENDYIDNPSLAL 60
Db      1 MEPPIPQSAFLTPNSVWVQPLDLSRMSHSLQHPLTLPIQVKTSHVENDYIDNPSLAL 60
Qy      61 TTGPKRTGAGAPLAPTPACDODVTHHMTSFGSRPSSVSSSSSTSDQRLDHHMAPPV 120
Db      61 TTGPKRTGAGAPLAPTPACDODVTHHMTSFGSRPSSVSSSSSTSDQRLDHHMAPPV 120
Qy      121 ADQASPRAVRIQPKVHHCQPLDLKGPVAPPDLKHFLCEACGCKCKECASPTLPSCW 180
Db      121 ADQASPRAVRIQPKVHHCQPLDLKGPVAPPDLKHFLCEACGCKCKECASPTLPSCW 180
Qy      181 VCNQECLSAQTLVNYGTCMCLVQGIIFYHCTNEDDEGSCADHPCSCSRNSCCARMSFMGA 240
Db      181 VCNQECLSAQTLVNYGTCMCLVQGIIFYHCTNEDDEGSCADHPCSCSRNSCCARMSFMGA 240
Qy      241 LSVLPLCLCYLPATGCVKLAQRGYDRLRRPGCRCKHTNSVICKAASGDAKTSRPDKPF 299
Db      241 LSVLPLCLCYLPATGCVKLAQRGYDRLRRPGCRCKHTNSVICKAASGDAKTSRPDKPF 299

```

```

RESULT 4
US-09-370-398-5
; Sequence 5, Application us/09370398
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Montgomery, Julie R.
; TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
; FILE REFERENCE: 28110/35878
; CURRENT APPLICATION NUMBER: us/09/370, 398
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-370-398-5

Query Match      93.8%; Score 1554.5; DB 17; Length 300;
Best Local Similarity 92.7%; Pred. No. 6.6e-117;
Matches 278; Conservative 9; Mismatches 12; Indels 1; Gaps 1;

Qy      1 MEPPIPQSA-PLTPNSVWVQPLDLSRMSHSLQHPLTLPIQVKTSHVENDYIDNPSLA 59
Db      1 MEPPIPQSA-PLTPNSVWVQPLDLSRMSHSLQHPLTLPIQVKTSHVENDYIDNPSLA 60
Qy      60 LTTGPKRTGAGAPLAPTPACDODVTHHMTSFGSRPSSVSSSSSTSDQRLDHHMAPP 119
Db      60 LTTGPKRTGAGAPLAPTPACDODVTHHMTSFGSRPSSVSSSSSTSDQRLDHHMAPP 120
Qy      120 VADQASPRAVRIQPKVHHCQPLDLKGPVAPPDLKHFLCEACGCKCKECASPTLPS 179
Db      120 VADQASPRAVRIQPKVHHCQPLDLKGPVAPPDLKHFLCEACGCKCKECASPTLPS 180

```

QY 180 WVCNOECICSAQTLVNYGTCMCLVQGIIFYHCTNEDDEGSCADHPCCSRSNCCARMSFMG 239
 |||
 Db 181 WVCNOECICSAQTLVNYGTCMCLVQGIIFYHCTNEDDEGSCADHPCCSRSNCCARMSFMG 240
 QY 240 ALSVPLPCLLCYLPATGCVKLAQRGYDRLRRPGCRCKHTNSVYICKAASGDAKTSRDPKPF 299
 |||
 Db 241 ALSVPLPCLLCYLPATGCVKLAQRGYDRLRRPGCRCKHTNSVYICKAASGDAKTSRDPKPF 300

RESULT 5
 US-10-083-215-5

; Sequence 5. Application US/10083215
 ; GENERAL INFORMATION:

; APPLICANT: Ballinger, Dennis G.
 ; APPLICANT: Montgomer, Julie R.
 ; TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
 ; FILE REFERENCE: 28110/35878
 ; CURRENT APPLICATION NUMBER: US/10/083,215
 ; CURRENT FILING DATE: 2002-02-26
 ; PRIOR APPLICATION NUMBER: US/09/370,398
 ; PRIOR FILING DATE: 1998-08-06
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: Patentln Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 300
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-10-083-215-5

Query Match
 Best Local Similarity 93.8%; Score 1554.5; DB 26; Length 300;
 Matches 278; Conservative 9; Mismatches 12; Indels 1; Gaps 1;

QY 1 MEPPIPQSA-PLTPNSVMVQPLDLSRMSHRLQHPILTIPIOVKTSHEVDYIDNPSLA 59
 |||
 Db 1 MEPPVQSSVVPVSSVMVQPLDLSRMSHRLQHPILTIPIOMKTSHEVDYIDNPSLA 60
 QY 60 LTTGPKRTGAPBELAPTPARCDDVTHMISFGSPSSVSSSSSTSDORLLDHNAAPP 119
 |||
 Db 61 PATGPRRPGGPELAPTPARCDDVTHMISFGSPSSVSSSSSTSDORLLDHNAAPP 120
 QY 120 VADQSPRAVRIOPKVHCQPLDLKGPAVPELDKHFLLCEACGCKCKECSAPRTLPSC 179
 |||
 Db 121 VADQSPRAVRIOPKVHCQPLDLKGPAVPELDKHFLLCEACGCKCKECSAPRTLPSC 180
 QY 180 WVCNOECICSAQTLVNYGTCMCLVQGIIFYHCTNEDDEGSCADHPCCSRSNCCARMSFMG 239
 |||
 Db 181 WVCNOECICSAQTLVNYGTCMCLVQGIIFYHCTNEDDEGSCADHPCCSRSNCCARMSFMG 240
 QY 240 ALSVPLPCLLCYLPATGCVKLAQRGYDRLRRPGCRCKHTNSVYICKAASGDAKTSRDPKPF 299
 |||
 Db 241 ALSVPLPCLLCYLPATGCVKLAQRGYDRLRRPGCRCKHTNSVYICKAASGDAKTSRDPKPF 300

RESULT 6
 US-10-090-190-5

; Sequence 5. Application US/10090190
 ; GENERAL INFORMATION:

; APPLICANT: Ballinger, Dennis G.
 ; APPLICANT: Montgomer, Julie R.
 ; TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
 ; FILE REFERENCE: 28110/35878
 ; CURRENT APPLICATION NUMBER: US/10/090,190
 ; CURRENT FILING DATE: 2002-02-04
 ; PRIOR APPLICATION NUMBER: US/09/370,398
 ; PRIOR FILING DATE: 1998-08-06
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: Patentln Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 300
 ; TYPE: PRT

; ORGANISM: Mus musculus
 US-10-090-190-5

Query Match
 Best Local Similarity 93.8%; Score 1554.5; DB 26; Length 300;
 Matches 278; Conservative 9; Mismatches 12; Indels 1; Gaps 1;

QY 1 MEPPIPQSA-PLTPNSVMVQPLDLSRMSHRLQHPILTIPIOVKTSHEVDYIDNPSLA 59
 |||
 Db 1 MEPPVQSSVVPVSSVMVQPLDLSRMSHRLQHPILTIPIOMKTSHEVDYIDNPSLA 60
 QY 60 LTTGPKRTGAPBELAPTPARCDDVTHMISFGSPSSVSSSSSTSDORLLDHNAAPP 119
 |||
 Db 61 PATGPRRPGGPELAPTPARCDDVTHMISFGSPSSVSSSSSTSDORLLDHNAAPP 120
 QY 120 VADQSPRAVRIOPKVHCQPLDLKGPAVPELDKHFLLCEACGCKCKECSAPRTLPSC 179
 |||
 Db 121 VADQSPRAVRIOPKVHCQPLDLKGPAVPELDKHFLLCEACGCKCKECSAPRTLPSC 180
 QY 180 WVCNOECICSAQTLVNYGTCMCLVQGIIFYHCTNEDDEGSCADHPCCSRSNCCARMSFMG 239
 |||
 Db 181 WVCNOECICSAQTLVNYGTCMCLVQGIIFYHCTNEDDEGSCADHPCCSRSNCCARMSFMG 240
 QY 240 ALSVPLPCLLCYLPATGCVKLAQRGYDRLRRPGCRCKHTNSVYICKAASGDAKTSRDPKPF 299
 |||
 Db 241 ALSVPLPCLLCYLPATGCVKLAQRGYDRLRRPGCRCKHTNSVYICKAASGDAKTSRDPKPF 300

RESULT 7
 PCT-US01-27760-819

; Sequence 819, Application PC/TUS0127760
 ; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
 ; FILE REFERENCE: 21272-115/785
 ; CURRENT APPLICATION NUMBER: PCT/US01/27760
 ; CURRENT FILING DATE: 2001-10-11
 ; PRIOR APPLICATION NUMBER: 09/687,527
 ; PRIOR FILING DATE: 2000-10-12
 ; NUMBER OF SEQ ID NOS: 892
 ; SOFTWARE: Custom
 ; SEQ ID NO 819
 ; LENGTH: 266
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US01-27760-819

Query Match
 Best Local Similarity 87.1%; Score 1444; DB 32; Length 266;
 Matches 259; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 35 LTIPIOVKTSHEVDYIDNPSLALTTGPKRTGAPBELAPTPARCDDVTHMISFG 94
 |||
 Db 2 LTIPIOVKTSHEVDYIDNPSLALTTGPKRTGAPBELAPTPARCDDVTHMISFG 61
 QY 95 RPSVSSSSSTSDORLLDHNAAPPVADQSPRAVRIOPKVHCQPLDLKGPAVPELDK 154
 |||
 Db 62 RPSVSSSSSTSDORLLDHNAAPPVADQSPRAVRIOPKVHCQPLDLKGPAVPELDK 121
 QY 155 HFLCEACGCKCKECSAPRTLPSCWVCNOECICSAQTLVNYGTCMCLVQGIIFYHCTNED 214
 |||
 Db 122 HFLCEACGCKCKECSAPRTLPSCWVCNOECICSAQTLVNYGTCMCLVQGIIFYHCTNED 181
 QY 215 DEGCADHPCCSRSNCCARMSFMGALSVPLPCLLCYLPATGCVKLAQRGYDRLRRPGCR 274
 |||
 Db 182 DEGCADHPCCSRSNCCARMSFMGALSVPLPCLLCYLPATGCVKLAQRGYDRLRRPGCR 241
 QY 275 CKHTNSVYICKAASGDAKTSRDPKPF 299
 |||
 Db 242 CKHTNSVYICKAASGDAKTSRDPKPF 266

	Matches	142;	Conservative	48;	Mismatches	87;	Indels	51;	Gaps
QY	6	POSALPLTPNSVWY--QPLIDS--RMSHSFLQHPILTLPIIDQYKTSHVENDYIDNPSLALT	61						
Db	3	PONQHSQSSSLVLDIOPSLDSRPRLDYEREIOPFTAILSLQIKAIRGSNEY-----T	54						
QY	62	TGPKRTGGAGPELAP-----TPARCQDQVTHHVISF-----	92						
Db	55	EGPSVYKRPAPRTAPTRAEQKEHETHEIIPINVNNNEHRTSHLGHAVLPSNARGPILSSS	114						
QY	93	--SGRPSSVSSSSSTSDQRLDLDHAAAP--PVADQASPRAYRIQPKVHCOPDLKLKPAAP	149						
Db	115	TSTGSAASGSSMSASSREGLLGRSPPTPRVYGHNSERAIRQPKOLYD--DLKG-SLK	171						
QY	150	PELDKHFLLCEACGKCKCEKACSPRTLPSVCWQCNQECLSAQTLVNYGTMCMLVQGIYH	209						
Db	172	EDLTQHKETIECGCGKCGCECTAPRTLPSCLCANRQCLCSASMSVEYGTMCMLVKGIFYH	211						
QY	210	CTNEDDEBSCADHPSCSCSRNSCCCAWSPFGALSLVYLPCLLCTLPLPATGCVKYLAKRGCDRLR	269						
Db	232	CSNDEGDSYSNPNPSCSQSHCCSRILCGAMSLFLPCLCYPPAKGCLKLCNRQCYDMTH	291						
QY	270	RPGCRCRHTNSVYICAAAS----GDKQTS	293						
Db	292	RPGCRCRKSNTYVYCKLDESCPSRQGGKPS	319						

```

RESULT 11
US-60-139-566-9
; Sequence 9, Application US/60139566
GENERAL INFORMATION:
APPLICANT: Yue, Henry
APPLICANT: Tang, Y. Tom
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Bandman, Olga
APPLICANT: Baughn, Mariah R.
APPLICANT: Azimzai, Yalda
APPLICANT: Yang, Junming
TITLE OF INVENTION: INTRACELLULAR SIGNALING MOLECULES
FILE REFERENCE: PF-0706 P
CURRENT APPLICATION NUMBER: US/60/139, 566
CURRENT FILING DATE: 1999-06-16
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PERL Program
SEQ ID NO 9
LENGTH: 319
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incycle Clone No: 2316805
PUBLICATION INFORMATION:
US-60-139-566-9

```

```

Query Match      41.0%: Score 679.5; DB 24; Length 319;
Best Local Similarity 43.3%: Pred. No. 3e-46;
Matches 142; Conservative 48; Mismatches 87; Indels 51; Gaps 9

QY 6 POSABLTGPNVWV--QPLDLS--RMSHSRLQPLRLTLPIDQVKTSHVEDYIDNPSTALT 61
   : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 PQNQGSSSLVVIQQLPULDSRQRLDYEREIQTPLALSDQLKATNGSEY-----T 54

QY 62 TGPKRTGCAPELAP-----TARCDQDVTNHWISF----- 92
   : : : : : : : : : : : : : : : : : : : : : : : :
Db 55 EGPVYVKRPAPPTARQOEKHEHRTHEITIPVANNNEHRTSHLGAVLPSMARGPILSR 114

QY 93 --SGRPSSVSSSTSSDQRLDHNHAP--PVADQASPRVRLQPKVNHQQLDLPKPAVP 149
   : : : : : : : : : : : : : : : : : : : : : : : :
Db 115 TSTGSAASSGSSNASSSEQGLGRSPPTPYVGHSERAKIRQPKQLYD--DLKG-SLK 171

QY 150 PELDKHFLLCERAGCKCKEKCASPTLPSGCWQCNQBLCSAOTLVNVTGCMCLVQGIYH 209

```

[illegible]

```

RESULT 12
US-10-018-170-9
; Sequence 9, Application US/10018170
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: LAL, Preeti
; APPLICANT: BADMAN, Olga
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: YANG, Junning
; APPLICANT: REDDY, Roopa
; APPLICANT: LU, Dyrong Alina M.
; TITLE OF INVENTION: INTRACELLULAR SIGNALING MOLECULES
; FILE REFERENCE: PR-0733 PCT
; CURRENT APPLICATION NUMBER: US/10/018.170
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 60/149,640; 60/164,417
; PRIOR FILING DATE: 1999-08-17; 1999-11-09
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: incyte ID No: 2316805CD1
US-10-018-170-9

```

```

Query Match Similarity      41.0%; Score 679.5; DB 26; Length 319;
Best Local Similarity      43.3%; Pred. No. 3e+6;
Matches 142; Conservative  48; Mismatches 87; Indels 51; Gaps 9;

Qy      6 POSAALTPNSVAV--QPLIDS--RMSHSRILPPLTLPIDQYKTSHVENDIDNPSLAT 61
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      3 PQNNGSSSLVLIQPLDLSRQRLDYEREIPTALISLDQIKALRGSNEY-----T 54

Qy      62 TGPKRTGCAPELAP-----TPARCDQVTHMVISF----- 92
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      55 EGPVSVKRPAPRTAPAROEKHERTHEIIPINANNNEHRTSHLGAVLPSNARGPILSKS 114

Qy      93 ---SGRPSSVSSSTSSNORLIDHMAPP--PVADQASPAVRVQRPVYHMQPLDGLGNAP 149
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      115 TSTGSAASSGSSSSASSSQGLGSRPPRPVGNHRSERAIROPQOLYD--DLKG-SLK 171

Qy      150 PELDHFLLIENCGKCKEKCASPTLPSPSCWNOBCLCAOTLVNNGTCMLVGGIFVH 209
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      172 EDLTDHKKRIKTCGCGKCKGECPTAPRTLRGLACNQCCLCSASMSVEYTGCMKVAGIFVH 231

Qy      210 CTNEDDEGSCADHPCSGSRNSNCCANMFWGALISVVLPCLLLYLPATGCVYKTLAQRGYDRLR 269
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      232 CSNDEGDSYSNPNPCSGSQSHCCSYTLCMGAMSLFLPCLCYPAKGCILKICRCRYDWIH 291

Qy      270 RRGCRCKHTNSVICKAAS---GDKKTS 293
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      292 RRGCRCKNSMTVYCKLESCPSRGGKPS 319

```

```

RESULT 13
US-09-488-725A-5405
; Sequence 5405, Application US/09488725A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/09/488,725A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7144
; SOFTWARE: PL_FL_genes_b Versions 1.0
; SEQ ID NO 5405
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-488-725A-5405

Query Match
Best Local Similarity 41.0%; Score 679.5; DB 18; Length 326;
Matches 142; Conservative 48; Mismatches 87; Indels 51; Gaps 9;

QY 6 POSAPLTPNSVYV--QPLDLS--RMSHSRLQHPITLPIIDQVKTSHVENDYIDNPSLALT 61
DB 10 PONOHGSGSSLVYIOQPSLDSRDLDEREIOPTAILSDQIKAIRGSNEY-----T 61
QY 62 TGPRTKGAPLAP-----TPARCDOVTHHWSF----- 92
DB 62 EGPVSVKRPARPTAPROEKHERTHEIIPINNNYEHRTSHLGHAVLPNAPGPIILRS 121
QY 93 --SGRPSSVSSSSSTSDORLLDHMAP--PVADQASPRAVRIQPVVHCQPLDLKGPAPV 149
DB 122 TSTGSAASSGSSASSSEGGILGRSPPTRPVPGHRSERAIPTQPKQLIVD--DLKG-SLK 178
QY 150 PELDKHFLICACGCKCKEBCASPTLPSGWCNOECLCSAOTLYNVTGCMCLVQGIFFH 209
DB 179 EDLTHKHFICGCKGCKGECTAPRTLPSCLAACNQCCLCSAESWVEYGTGCMCLVQGIFFH 238
QY 210 CTNEDDEGSCADHPGCSGSRNCCARMSFMGALSVVLPCLLCYLPATGCVKLAQRCYDRLR 269
DB 239 CSNDEGDSYSDNPCSCGSHCCSRYLGMGAMSLFLPCLLCYLPAPKAGCLICRCRCYDWH 298
QY 270 RPGCRCKHNTSVYICKAAS----GDAKTS 293
DB 299 RPGCRCKNSNTVYCKLESCPSRGGKPS 326

```

```

; CURRENT FILING DATE: 1999-09-10
; EARLIER APPLICATION NUMBER: US 60/099,936
; EARLIER FILING DATE: 1998-09-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US99-20638-7

Query Match
Best Local Similarity 41.0%; Score 679.5; DB 1; Length 338;
Matches 142; Conservative 48; Mismatches 87; Indels 51; Gaps 9;

QY 6 POSAPLTPNSVYV--QPLDLS--RMSHSRLQHPITLPIIDQVKTSHVENDYIDNPSLALT 61
DB 22 PONOHGSGSSLVYIOQPSLDSRDLDEREIOPTAILSDQIKAIRGSNEY-----T 73
QY 62 TGPRTKGAPLAP-----TPARCDOVTHHWSF----- 92
DB 74 EGPVSVKRPARPTAPROEKHERTHEIIPINNNYEHRTSHLGHAVLPNAPGPIILRS 133
QY 93 --SGRPSSVSSSSSTSDORLLDHMAP--PVADQASPRAVRIQPVVHCQPLDLKGPAPV 149
DB 134 TSTGSAASSGSSASSSEGGILGRSPPTRPVPGHRSERAIPTQPKQLIVD--DLKG-SLK 190
QY 150 PELDKHFLICACGCKCKEBCASPTLPSGWCNOECLCSAOTLYNVTGCMCLVQGIFFH 209
DB 191 EDLTHKHFICGCKGCKGECTAPRTLPSCLAACNQCCLCSAESWVEYGTGCMCLVQGIFFH 250
QY 210 CTNEDDEGSCADHPGCSGSRNCCARMSFMGALSVVLPCLLCYLPATGCVKLAQRCYDRLR 269
DB 251 CSNDEGDSYSDNPCSCGSHCCSRYLGMGAMSLFLPCLLCYLPAPKAGCLICRCRCYDWH 310
QY 270 RPGCRCKHNTSVYICKAAS----GDAKTS 293
DB 311 RPGCRCKNSNTVYCKLESCPSRGGKPS 338

RESULT 15
US-09-393-314B-7
; Sequence 7, Application US/09393314B
; GENERAL INFORMATION:
; APPLICANT: EDMONDS, Brian T.
; TITLE OF INVENTION: ANTAGONISTS OF FIBROBLAST GROWTH FACTOR
; FILE REFERENCE: X-12136
; CURRENT APPLICATION NUMBER: US/09/393,314B
; CURRENT FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: US 60/099,936
; PRIOR FILING DATE: 1998-09-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-393-314B-7

Query Match
Best Local Similarity 41.0%; Score 679.5; DB 17; Length 338;
Matches 142; Conservative 48; Mismatches 87; Indels 51; Gaps 9;

QY 6 POSAPLTPNSVYV--QPLDLS--RMSHSRLQHPITLPIIDQVKTSHVENDYIDNPSLALT 61
DB 22 PONOHGSGSSLVYIOQPSLDSRDLDEREIOPTAILSDQIKAIRGSNEY-----T 73
QY 62 TGPRTKGAPLAP-----TPARCDOVTHHWSF----- 92
DB 74 EGPVSVKRPARPTAPROEKHERTHEIIPINNNYEHRTSHLGHAVLPNAPGPIILRS 133

```

Sat May 11 12:51:24 2002

us-10-082-902-2.ram

Page 7

[illegible]

Search completed: May 8, 2002, 10:48:30
Job time: 162 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:45:48 ; Search time 12.56 Seconds

(Without alignments)
535,707 Million cell updates/sec

Title: US-10-082-902-2

Perfect score: 1657
Sequence: 1 MEPP1PQSAFLFNSWVWP.....SVICKAASGDATSRDPKPF 299

Scoring table: BLOSUM62

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfills1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	650	39.2	117 3	US-08-965-903B-16 Sequence 16, Appl
2	641.5	38.7	315 3	US-08-965-903B-8 Sequence 8, Appl
3	528	31.9	139 3	US-08-965-903B-20 Sequence 20, Appl
4	503.5	30.4	116 3	US-08-965-903B-9 Sequence 9, Appl
5	493	29.8	117 3	US-08-965-903B-6 Sequence 6, Appl
6	396	23.9	591 3	US-08-965-903B-2 Sequence 2, Appl
7	368.5	22.2	124 3	US-08-965-903B-4 Sequence 4, Appl
8	301	18.2	83 3	US-08-965-903B-13 Sequence 13, Appl
9	237	14.3	100 3	US-08-965-903B-11 Sequence 11, Appl
10	214	12.9	45 3	US-08-965-903B-18 Sequence 18, Appl
11	202	12.0	45 3	US-08-965-903B-17 Sequence 17, Appl
12	199	12.0	45 3	US-08-965-903B-19 Sequence 19, Appl
13	120	7.2	2476 2	US-08-276-967-2 Sequence 2, Appl
14	106	6.4	721 3	US-08-872-855-7 Sequence 7, Appl
15	103.5	6.2	2050 2	US-08-347-594A-2 Sequence 2, Appl
16	103.5	6.2	2050 2	US-08-347-594A-2 Sequence 2, Appl
17	102.5	6.2	577 2	US-07-728-215-29 Sequence 29, Appl
18	99.5	6.0	578 2	US-08-981-392-13 Sequence 13, Appl
19	98	5.9	2556 1	US-08-185-432-17 Sequence 17, Appl
20	98	5.9	2556 1	US-08-083-590A-20 Sequence 20, Appl
21	96.5	5.8	3111 2	US-08-533-384-20 Sequence 20, Appl
22	96.5	5.8	3111 2	US-08-460-309-4 Sequence 4, Appl
23	96.5	5.8	3111 2	US-08-125-077-4 Sequence 4, Appl
24	95.5	5.8	2471 1	US-08-185-432-16 Sequence 16, Appl
25	95.5	5.8	2471 1	US-08-083-580A-19 Sequence 19, Appl
26	95.5	5.8	2471 1	US-08-532-384-19 Sequence 19, Appl
27	95	5.7	676 3	US-08-630-172-10 Sequence 10, Appl

28	95	5.7	676	4	US-09-375-419-10	Sequence 10, Appl
29	95	5.7	769	2	US-08-789-078-1	Sequence 1, Appl
30	95	5.7	769	2	US-08-752-633-1	Sequence 1, Appl
31	95	5.7	769	2	US-08-476-062A-45	Sequence 45, Appl
32	95	5.7	769	2	US-07-728-215-31	Sequence 31, Appl
33	95	5.7	769	5	PCT-US95-04886-1	Sequence 1, Appl
34	95	5.7	769	5	PCT-US96-01314-45	Sequence 45, Appl
35	95	5.7	1219	4	US-08-882-046-5	Sequence 5, Appl
36	95	5.7	2813	3	US-08-896-449A-2	Sequence 2, Appl
37	95	5.7	2813	3	US-09-132-652-2	Sequence 2, Appl
38	94.5	5.7	1833	3	US-08-479-722B-2	Sequence 2, Appl
39	94.5	5.7	1833	3	PCT-US95-02251-18	Sequence 18, Appl
40	94	5.7	314	2	US-08-460-309-19	Sequence 19, Appl
41	94	5.7	314	2	US-08-125-077-19	Sequence 7, Appl
42	94	5.7	2594	4	US-08-718-388-7	Sequence 9, Appl
43	94	5.7	5405	4	US-08-718-388-9	Sequence 9, Appl
44	93.5	5.6	410	6	5177197-1	Patent No. 5177197
45	93.5	5.6	1394	6	5177197-30	Patent No. 5177197

ALIGNMENTS

RESULT 1
US-08-965-903B-16
Sequence 16, Application US/08965903B
Patent No. 6060275
GENERAL INFORMATION:
APPLICANT: Hachon, Nir
TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
TITLE OF INVENTION: SEQUENCE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESS: Dellinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 60/030232
FILING DATE: 07-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Pelthor, Joanne R
REGISTRATION NUMBER: 42,995
REFERENCE/DOCKET NUMBER: 8600-0177.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
TELEX:
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: m-spry4 cysteine-rich region
US-08-965-903B-16

Query Match 39.2%: Score 650; DB 3; Length 117;
Best Local Similarity 97.4%; Pred. No. 1.1e-51;
Matches 114; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 183 NOECLSAQTLVNYGTCMCLVGIHYHCTNEDDEGSCADHPSCSRSNCCARMSFGALS 242
DB 1 NOECLSAQTLVNYGTCMCLVGIHYHCTNEDDEGSCADHPSCSRSNCCARMSFGALS 60
OY 243 VVPLCLCYLPATGCVKLAQRGIDRLRRPGCRCKHTNSVICKAASGDTKTSRDKPF 239
DB 61 VVPLCLCYLPATGCVKLAQRGIDRLRRPGCRCKHTNSVICKAASGDTKTSRDKPF 117

RESULT 2

US-08-965-903B-8
Sequence 8, Application US/08965903B
Patent No. 6060275
GENERAL INFORMATION:
APPLICANT: Hachow, Nir
TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,903B
FILING DATE: 07-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030232
FILING DATE: 07-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Petithory, Joanne R
REGISTRATION NUMBER: 42,995
REFERENCE/DOCKET NUMBER: 8600-0177.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 315 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: h-spry2 protein
US-08-965-903B-8

Query Match 38.7%: Score 641.5; DB 3; Length 315;
Best Local Similarity 43.6%; Pred. No. 2.3e-50;
Matches 130; Conservative 45; Mismatches 84; Indels 39; Gaps 8;

OY 19 QPLDSRMSHR-----LOHPLTLPIDQVKTSHENDYIDNPSIALTTGPK--- 65
DB 12 QPLQTRDGRGGRGEPDRALNQVHVLSDIRAIRNNTETDEPTVPRGLKAP 71
OY 66 -----RTTGAPDELAPP-----ARCDQDVTHHWSFSCRPSVSSSSSTSS 107

DB 72 RPSQHKHERHLGLPEHRPPRLQHSQVHSSARAPLSRSISTVSSGSSRSTRTSTSSSS 131
OY 108 DQGLL-DHMAPPVADQAPRAVRIOPKVYHGOPLDKGPAPELDKFLCEAGCKK 166
DB 132 EORLLSSFSFGVAD----GIRVOPK--SELKPELK-PLSKEDGLHATYCEDCGCK 185
OY 167 CKRCASPTPLPSCWVNOECLSAQTLVNYGTCMCLVGIHYHCTNEDDEGSCADHPSC 226
DB 186 CKRCCTPRPLPSWIDKQCLSAQNVIDYGTCCVKGLFPHCSN-DEEDCADHPSC 244
OY 227 SRNCCARMSFGALS VVPLCLCYLPATGCVKLAQRGIDRLRRPGCRCKHTNSVICK 284
DB 245 SQSHCCTRWSAMGVMSLFLPCLMICYLPAGCKLCKGQCYDVRNRPGRCKNSNTVCK 302

RESULT 3

US-08-965-903B-20
Sequence 20, Application US/08965903B
Patent No. 6060275
GENERAL INFORMATION:
APPLICANT: Hachow, Nir
TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,903B
FILING DATE: 07-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030232
FILING DATE: 07-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Petithory, Joanne R
REGISTRATION NUMBER: 42,995
REFERENCE/DOCKET NUMBER: 8600-0177.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
TELEX:
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 139 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: h-spry1
US-08-965-903B-20

Query Match 31.9%: Score 528; DB 3; Length 139;
Best Local Similarity 62.6%; Pred. No. 1.4e-40;
Matches 87; Conservative 24; Mismatches 24; Indels 4; Gaps 1;

OY 159 CEACGCKCKECSPTPLPSCWVNOECLSAQTLVNYGTCMCLVGIHYHCTNEDDEGS 218
DB 1 CEACGCKCKECSPTPLPSCWVNOECLSAQTLVNYGTCMCLVGIHYHCTNEDDEGS 60

OY	219	CADHSCSRSNCARNSFMALSVLPCLLCYYPAGCYKLAORGDLRHPGCRCKHM	278
		: : : : : : : : : : : : : :	
Dd	61	YSDFPCSSQSHCCSRYLCKGMAMSLFLPCLLCYPAGCGLKLGRCTDWMTHRGCRCKNS	120
OY	279	NSVYCKAAS-----DATKS	293
		:	
Dd	121	NIVYCKLESFCPSRGCGKPS	139

RESULT 4
US-08-965-903B-9
; Sequence 9, Application US/08965903B
; Patent No. 6060375

```

1  GENERAL INFORMATION:
2  APPLICANT:  Hacohen, Nir
3  APPLICANT:  Krasnow, Mark A.
4  TITLE OF INVENTION:  SPOUTRY PROTEIN AND CODING
5  TITLE OF INVENTION:  SEQUENCE
6  NUMBER OF SEQUENCES:  20
7  CORRESPONDENCE ADDRESS:

```

STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965, 903B
FILING DATE: 07-NOV-1997

INDIVIDUAL ISOLATE: h-spr2 cysteine-rich region
US-08-965-9038-9

Query Match	30.4%	Score 503.5;	DB 3;	Length 116;
Best Local Similarity	67.5%	Pred. No. 1.9e-38;		
Matches 79;	Conservative 19;	Mismatches 18;	Indels 1;	Gaps 1

QY 159 CEAAGCKKCKEACASPTLPSCWVNOGLCSAQTLYNNGKMLVGGFYHYHTNDDGGS 218
 1 CEECGKCKCKEETRYPLPSMDICDKQLCSAQNADIDYGTCCVCKGGLFYHCNSN-DEEDN 59
 QY 219 GADHPSCSCRSNCKARMSFMAKLSVLPCLLTLYLPATGCVKLAQKGRDRLRRPGCRC 275
 60 GADNPSCSQSHSCCTRMASAMKVMSLFPLPLMCTLYLPARKGKJLKLDGGYVRYVNPSPGRC 116

; Sequence 6, Application US/08965903B
; Patent No. 6060275

1 APPLICANT: Kacohen, Nir
 2 APPLICANT: Krasnow, Mark A.
 3 TITLE OF INVENTION: SPOURRY PROTEIN AND CODING
 4 TITLE OF INVENTION: SEQUENCE
 5 NUMBER OF SEQUENCES: 20
 6 CORRESPONDENCE ADDRESS:
 7 ADDRESSEE: Dehlinger & Associates
 8 STREET: 350 Cambridge Ave., Suite 250
 9 CITY: Palo Alto
 10 STATE:

COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,903B
FILING DATE: 07-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030332

INDIVIDUAL ISOLATE: h-spry1 cysteine-rich region
US-08-965-903B-6

Query Match	29.8%	Score 493;	DB 3;	Length 117;
Best Local Similarity	66.7%	Pred. No. 1.76-37;		
Matches	78;	Conservative	21;	Mismatches 18;
				Indels 0;
				Gaps 0;

QY 159 CAGACGCKCKEACASRTPLSPCAWNOEELASACTLAVYGCMMKLVGGIYHCTNNDEBGS 210
Db 1 CEGCGCKCKGECTARTPLPCLACACROCLCSABEMETGYCCKLVAGIYHCSNDBEDGS 60
QY 219 CADHPSCGSRSMCCARMSFMGALSVLPCLLCYLPATGCVKTLAQRGYDLRRPAGCRC 275
Db 61 YSNPSCGSCSHSCGSRYLCTMGAMSLFPLCLCYCPAPAGGLCTKLTRRCYDIDINHPGRC 117

RESULT 6
 US-08-965-903B-2
 : Sequence 2, Application US/08965903B
 : Patent No. 6060275
 : GENERAL INFORMATION:
 : APPLICANT: Hachosen, Nlr
 : APPLICANT: Krasnow, Mark A.
 : TITLE OF INVENTION: SPOURY PROTEIN AND CODING
 : TITLE OF INVENTION: SEQUENCE
 : NUMBER OF SEQUENCES: 20
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Dehlinger & Associates
 : STREET: 350 Cambridge Ave., Suite 250

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,903B
FILING DATE: 07-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030232
FILING DATE: 07-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Pelthorj, Joanne R
REGISTRATION NUMBER: 42,995
REFERENCE/DOCKET NUMBER: 8600-0177.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 591 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-965-903B-2

Query Match 23.9%; Score 396; DB 3; Length 591;
Best Local Similarity 29.4%; Pred. No. 7.3e-28;
Matches 95; Conservative 50; Mismatches 120; Indels 58; Gaps 10;

QY 4 PIPGAPLTPNSVWVQPLDLSMSH-SRLQHPRLTLPIDQVTSHEVDYIDNPSLATLT 62
DB 190 PREPERLITNEYV-----DPLQHATRSQHPAGQDNGOTTTHLLLPBNOHLLHQ 243
QY 63 GKRTGAGAPLAPTPARCDQVTHHWISFGSPSSVSSSSSTSD-----ORLDDM 115
DB 244 HQHQLQQQQQQQQQQQQQQHQLHQHQQHARLATTQATSVSDHTDGLHSHLQST 303
QY 116 APPPADQ-ASPR-----AVRIQKVYHCQPLDK----- 144
DB 304 TKPRASKQPALPRLGMLGIGLGLNLPITTKQPTPAOKERHMALELLQPGGAGNG 363
QY 145 GRAV-----PRLDKHFLLEACGKCKKCEKASPTLPSCWVNOECLESAOTLVNYGTCM 200
DB 364 GPLVWAGPSSLNP--LYCPGRCRCRCQCSPPRLPQTWCKNTCLCSASVVDYASCL 421
QY 201 CLVQGIYHCTNED---DEGS---CADHPGCSHSNCCARMSFGALSVVLPCLLCYLP 253
DB 422 CCAKALFYHCARDNDLDDDDGNGTGPCVDNPGSCGYKRTQRMGMGLSIFLPCIMFYWP 481
QY 254 ATGCYKLAORGYDRLRRGCRCK 276
DB 482 MRGCKLCEKCYGRFAGRCRCQ 504

RESULT 7
US-08-965-903B-4
Sequence 4, Application US/08965903B
Patent No. 6060275
GENERAL INFORMATION:
APPLICANT: Hachoen, NLR
APPLICANT: Krasnow, Mark A.
TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
TITLE OF INVENTION: SEQUENCE

NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,903B
FILING DATE: 07-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030232
FILING DATE: 07-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Pelthorj, Joanne R
REGISTRATION NUMBER: 42,995
REFERENCE/DOCKET NUMBER: 8600-0177.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 124 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: d-spry cysteine-rich region
US-08-965-903B-4

Query Match 22.2%; Score 368.5; DB 3; Length 124;
Best Local Similarity 46.0%; Pred. No. 3.2e-26;
Matches 57; Conservative 25; Mismatches 35; Indels 7; Gaps 2;

QY 159 CEACGKCKKCEKASPTLPSCWVNOECLESAOTLVNYGTCMCLVQGIYHCTNED----- 214
DB 1 CPGRCRCRCQCSPPRLPQTWCKNTCLCSASVVDYASCLCAKALFYHCARDNDLDC 60
QY 215 DEGS---CADHPGCSHSNCCARMSFGALSVVLPCLLCYLPATGCVKLAORGYDRLRRP 271
DB 61 DDGNGTGPCVDNPGSCGYKRTQRMGMGLSIFLPCIMFYWPARGCKLCEKCYGRFAGR 120
QY 272 GCRC 275
DB 121 GCRC 124

RESULT 8
US-08-965-903B-13
Sequence 13, Application US/08965903B
Patent No. 6060275
GENERAL INFORMATION:
APPLICANT: Hachoen, NLR
APPLICANT: Krasnow, Mark A.
TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA

```

1 CLASSIFICATION: 435
2 PRIOR APPLICATION DATA:
3 APPLICATION NUMBER: US 60/030322
4 FILING DATE: 07-NOV-1996
5 ATTORNEY/AGENT INFORMATION:
6 NAME: Pettithory, Joanne R
7 REGISTRATION NUMBER: 42,995
8 REFERENCE/DOCKET NUMBER: 8600-0177.30
9 TELECOMMUNICATION INFORMATION:
10 TELEPHONE: 650-324-0880
11 TELEFAX: 650-324-0960
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
```

TELEFAX: 650-324-0960
TELEX:
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: h-spry1 highly conserved cysteine-rich region
US-08-965-903B-18

Query Match 12.9%; Score 214; DB 3; Length 45;
Best Local Similarity 73.3%; Pred. No. 8e-13;
Matches 33; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

OY 159 CEACGKCKCKECASPTPLPSCWCNQCICSAOTLVNYGTMCIV 203
Db 1 CEACGKCKCKECCTAPRLPCLACNRCICSAESMVEYGTMCIV 45

RESULT 11
US-08-965-903B-17
Sequence 17, Application US/08965903B
Patent No. 6060275
GENERAL INFORMATION:
APPLICANT: Hachon, Nir
TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,903B
FILING DATE: 07-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030232
FILING DATE: 07-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Pelithory, Joanne R
REGISTRATION NUMBER: 42,995
REFERENCE/DOCKET NUMBER: 8600-0177.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
TELEX:
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: h-spry2 highly conserved cysteine-rich region
US-08-965-903B-17

Query Match 12.2%; Score 202; DB 3; Length 45;

Best Local Similarity 66.7%; Pred. No. 9.7e-12;
Matches 30; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

OY 159 CEACGKCKCKECASPTPLPSCWCNQCICSAOTLVNYGTMCIV 203
Db 1 CEACGKCKCKECCTAPRLPCLACNRCICSAESMVEYGTMCIV 45

RESULT 12
US-08-965-903B-19
Sequence 19, Application US/08965903B
Patent No. 6060275

GENERAL INFORMATION:
APPLICANT: Hachon, Nir
TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,903B
FILING DATE: 07-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030232
FILING DATE: 07-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Pelithory, Joanne R
REGISTRATION NUMBER: 42,995
REFERENCE/DOCKET NUMBER: 8600-0177.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
TELEX:

INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: h-spry3 highly conserved cysteine-rich region
US-08-965-903B-19

Query Match 12.0%; Score 199; DB 3; Length 45;
Best Local Similarity 64.4%; Pred. No. 1.8e-11;
Matches 29; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

OY 159 CEACGKCKCKECASPTPLPSCWCNQCICSAOTLVNYGTMCIV 203
Db 1 CEACGKCKCKECCTAPRLPCLACNRCICSAESMVEYGTMCIV 45

RESULT 13
US-08-276-967-2
Sequence 2, Application US/08276967
Patent No. 5851817
GENERAL INFORMATION:
APPLICANT: Hardy, Daniel M.
APPLICANT: Garders, David L.

TITLE OF INVENTION: Species-Specific Egg-Binding Proteins of
TITLE OF INVENTION: Sperm
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,967
FILING DATE: Submitted Herewith
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: UTSD:418\KIT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-787-1400
TELEFAX: 713-789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2476 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-276-967-2

Query Match 7.2%; Score 120; DB 2; Length 2476;
Best Local Similarity 19.8%; Pred. No. 0.036;
Matches 63; Conservative 33; Mismatches 112; Indels 110; Gaps 17;

QY 2 EPIPOSALVTPNSVWVQPLDSRMSHRLQHP--TILPIOVKTSHEVNDYIDNPSLA 59
DB 552 KPIVPEKSVPEKPTVTEETITETKLYPTERTTPTTKRTTPTRT---TTPTR 608
QY 60 LTGPRTRGAPBELAPPARCDQVTHMISFSGRPSVSSSSSTSSDRLDHPAPP 119
DB 609 TTPTRT-----TTPTRTPTPTPTT-TTTPKKTTPTEKTIITPTERTI-----AP 654
QY 120 VADQASPRVRIQPKVY-----HCQPLDLKGPVAPPRLDHPFLCEAC 162
DB 655 TTPQSPPTLVTPQAPVAVMPTSATVTPRTTASCP-----NAHF----- 696
QY 163 GKCKCK-ECASPTPLPSCVW-CNOELC-----SAQTLVNYGTGCMCLVQGFYHCTNE 214
DB 697 ERACAGVSCSP--TPNCLFPCKPGCVCDPGLFSGSHCVNASSCCDFYNDNTYKLGTDW 754
QY 215 DEGSADHPSC--SRNC-----CA----- 233
DB 755 FSPNCTEH-CHGRPSRMCQTFKCGTHVQQLKNQGYGCHPYGATVCSVYGDPHYLTFD 813
QY 234 --RMSFMGLSVLV--PC 247
DB 814 GRFENFMKCTYLAQPC 831

RESULT 14
US-08-872-855-7
Sequence 7, Application US/08872855
Patent No. 6121045
GENERAL INFORMATION:
APPLICANT: McCarthy, Sean
APPLICANT: Gearing, David

TITLE OF INVENTION: NOVEL HUMAN DELTA3 COMPOSITIONS AND
TITLE OF INVENTION: THERAPEUTIC USES THEREFOR
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/872,855
FILING DATE: 11-JUN-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MAA-003.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 721 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-872-855-7

Query Match 6.4%; Score 106; DB 3; Length 721;
Best Local Similarity 24.2%; Pred. No. 0.14;
Matches 55; Conservative 13; Mismatches 79; Indels 80; Gaps 14;

QY 138 COP-----LDLKPAPVPELDKHFLLCEAGCKCK-----ECASPTPLPSC----- 179
DB 214 CNPKGKGLXCTPTICLPDCDEHNGYCDKPRGECKVGVGQGRICDCT---RTPGLMOTC 270
QY 180 ---WVCN-OE-----CLC-----SAQTLVNYG-----TCMCLVQGFYHCTNE 213
DB 271 QQPQCNCOEGGGLFCQNDLNYCTHMKPCENGATCTMTGGGYSYCSGRPGYTSNCEIE 330
QY 214 DDE-----GSCAD-----HPCSC-----SRNCCARMSFMGL-----S 242
DB 331 VNECDANPKNGSCSDLENSYTCSPGFGYKNCLELSAMTCADGPCFNGRCADNPDG 390
QY 243 VLPCLCYLPATGCVKLAQRGYRLRRPGCRCKHT-NSVICKAASG 288
DB 391 YICPVPVGT-SGFNCEKRIIDYCNSNFCANGARCELDGNSYICQCOEG 436

RESULT 15
US-08-961-392-5
Sequence 5, Application US/08961392
Patent No. 6262025
GENERAL INFORMATION:
APPLICANT: Ish-Horowitz, David
APPLICANT: Henrique, Domingos Manuel Pinto
APPLICANT: Lewis, Julian Hart
APPLICANT: Artavanis-Tsakonas, Spyridon
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES
TITLE OF INVENTION: OF VERTEBRATE DELTA GENES AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:45:43 ; Search time 23.43 Seconds
(without alignments)

(without alignments)
945.280 Million cell updates/sec

Title: US-10-082-902-2
perfect score. 1657

Sequence: 1 MEPIPIQSA¹LT²PN³SV⁴MV⁵Q⁶.....SVICKASGDAKTSRPDKPF 299

Scoring table:	BLOSUM62	Gapopen 10	Gapext 0.5
----------------	----------	------------	------------

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
Maximum Watch	100%	

Database :

1:	/S1D8_gcgdata/gemseq/gemseqp/AA1980.DAT *
2:	/S1D8_gcgdata/gemseq/gemseqp/AA1981.DAT *
3:	/S1D8_gcgdata/gemseq/gemseqp/AA1982.DAT *
4:	/S1D8_gcgdata/gemseq/gemseqp/AA1983.DAT *
5:	/S1D8_gcgdata/gemseq/gemseqp/AA1984.DAT *
6:	/S1D8_gcgdata/gemseq/gemseqp/AA1985.DAT *
7:	/S1D8_gcgdata/gemseq/gemseqp/AA1986.DAT *
8:	/S1D8_gcgdata/gemseq/gemseqp/AA1987.DAT *
9:	/S1D8_gcgdata/gemseq/gemseqp/AA1988.DAT *
10:	/S1D8_gcgdata/gemseq/gemseqp/AA1989.DAT *
11:	/S1D8_gcgdata/gemseq/gemseqp/AA1990.DAT *
12:	/S1D8_gcgdata/gemseq/gemseqp/AA1991.DAT *
13:	/S1D8_gcgdata/gemseq/gemseqp/AA1992.DAT *
14:	/S1D8_gcgdata/gemseq/gemseqp/AA1993.DAT *
15:	/S1D8_gcgdata/gemseq/gemseqp/AA1994.DAT *
16:	/S1D8_gcgdata/gemseq/gemseqp/AA1995.DAT *
17:	/S1D8_gcgdata/gemseq/gemseqp/AA1996.DAT *
18:	/S1D8_gcgdata/gemseq/gemseqp/AA1997.DAT *
19:	/S1D8_gcgdata/gemseq/gemseqp/AA1998.DAT *
20:	/S1D8_gcgdata/gemseq/gemseqp/AA1999.DAT *
21:	/S1D8_gcgdata/gemseq/gemseqp/AA2000.DAT *
22:	/S1D8_gcgdata/gemseq/gemseqp/AA2001.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query Match	Length	ID	Description
No.					
1	679.5	41.0	319	22	Human polypeptide
2	679.5	41.0	319	22	Amino acid sequenc
3	679.5	41.0	326	22	Human polypeptide
4	672.5	40.6	319	21	Human SPROUTY-1 pr
5	668.5	40.3	338	21	Human SPROUTY-1 pr
6	654	39.5	294	21	Human SPROUTY-1 pr
7	653.5	39.4	288	21	Human SPROUTY-1 pr
8	653.5	39.4	295	21	Human SPROUTY-1 pr
9	641.5	38.7	315	19	Homo sapiens sprow
10	557.5	33.6	167	21	Human SPROUTY-1 pr
11	543.5	32.8	156	21	Human SPROUTY-1 pr

15	104.5	6.3	361	7	AAE0438
14	105	6.3	1554	20	AAE0430
43	105	6.3	1534	20	AAE0413
42	105	6.3	1554	20	AAE0410
41	105	6.3	1534	19	AAE4656
40	105	6.3	1508	20	AAE9670
39	105	6.3	1508	20	AAE0413
38	105	6.3	1508	20	AAE2714
37	105	6.3	721	21	AAE7903
36	105	6.3	353	8	AAE70437
35	107	6.5	602	21	AAE9566
34	107	6.5	582	21	AAE5944
33	107	6.5	561	21	AAE5944
32	107	6.5	561	21	AAE5944
31	108	6.5	1472	13	AAE2946
30	109	6.6	1784	21	AAE4137
29	109	6.6	1686	22	AAE7494
28	109	6.6	494	21	AAE2557
27	109	6.6	459	21	AAE2558
26	110.5	6.7	2813	7	AAE60462
25	110.5	6.8	2813	7	AAE60055
24	113.5	6.8	333	21	AAE32411
23	117	7.1	567	20	AAE26773
22	120	7.2	2476	20	AAE65793
21	121	7.3	114	21	AAE8233
20	126	7.6	738	20	AAE39552
19	142.5	8.6	250	21	AAE8333
18	237	14.3	100	19	AAAE45011
17	291	17.6	83	21	AAAE5051
16	291	17.6	83	21	AAAE45011
15	396	23.9	591	19	AAAE4793
14	443.5	26.8	140	19	AAAE4793
13	493	29.8	117	21	AAAE4189
12	528	31.9	139	21	AAAE4189

ALIGNMENTS

RESULT	1
AAAM38688	
ID	AAAM38688 standard; Protein; 319 AA.
XX	
AC	AAAM38688;
XX	
DT	22-OCT-2001 (first entry)
XX	
DE	Human polypeptide SEQ ID NO 1833.
XX	
KW	Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW	peripheral nervous system; neuropathy; central nervous system; CNS;
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW	leukaemia.
XX	
XX	
OS	Homo sapiens.
XX	
PN	MO20015312-A1.
PD	
XX	
XX	26-JUL-2001.
PF	
XX	
PF	26-DEC-2000; 2000MO-US34263.
XX	
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0598042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.
XX	

(HYSE-) HYSEQ INC.

PA (HXSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang
PI Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J,
PI Zhao QA, Zhou P, Goodrich R, Dirmnac RT;
XX
XX WPI: 2001-442253/47.
DR N-PSDB; AA157844.

PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -

Example 3; SEQ ID NO 1833; 10078pp; English.

CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AA836642-AA442213) with neurotrophic,
CC immunosuppressant and cytoskeletal activity. The polynucleotides
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

SQ Sequence 319 AA;

Query Match	Score	DB	Length
41.08;	679.5;	DB 22;	Length 319;

Matches 142; Conservative 48; Mismatches 87; Indels 51; Gaps 9,

```

QY 6 POSAFLTNSVWV---QPLTDS--RMSHSRLOHPTLTIIDOVKTSHEVNDIDNPSLALT 61
Db 3 pqnqngsgssalvngqprldasrprldyerelqpalalsldqikafrsney-----t 54
QY 62 TGPKRTKCGAEPLAR-----TPACDODVTNHWISF----- 92
Db 55 epgsvvxxpraprtbrpqkherthelalpinnnnnyehrhghavlpnsargpilsrs 114
QY 93 --SGRPSSVSSSSSTSSQOPLIDMAP--PVAODASPRAYVIOPKVYHCOPDLKGPAPV 149
Db 115 tsfgsaasngsnaasaseeqglilgtsprtprrpygrnrselarttqpkqlivd--dlkg-slk 171
QY 150 PELDKHFLCACGCKCKCEKASRPTLPSCWVCNOECLCSAQTLVNTGCMCLVOGIFTH 209
Db 172 edltqhkflceqgckckcgcetaprltlpsclacnrcrgclsaasmevgtcmclvtkgflth 231
QY 210 CTNEDDEGSCADHFCSCGRSNCRCARMSFGMALSVLCECLCYEPATCGCVKLAQDGRYRLR 269
Db 232 csnddegdsydsndpcoscsghncserylcmgamaflfpollcyprakgclnrttergywh 291
QY 270 RRGCRKRTNSVICKAAS----GIAKTS 293
Db 292 tpgtgcckmsnlvcklescpargtqgkps 319

```

RESULT 2

ID AAB64377 standard; Protein; 319 AA.

AC AAB64377;

DT 22-MAR-2001 (first entry)

DE Amino acid sequence of human intracellular signalling molecule INTRAG
XX

KM Human intracellular signaling molecule; *INTRA*; immunosuppressive;
KM cytosolic; neuroprotective; neurotropic; antiarteriosclerotic; cancer;
KM antiinflammatory; anti-HIV; neuroleptic; antiarterial; antifungal;
KM antiviral; antiparasitic; antihelminthic; antiparkinsonian; AIDS;
KM cell proliferative disorder; arteriosclerosis atherosclerosis; epilepsy;
KM inflammatory disorder; Addison's disease; gastrointestinal disorder;
KM neurological disorder; Parkinson's disease; Creutzfeldt-Jakob disease;
KM mental disorder; schizophrenia; anxiety.

OS Homo sapiens

PN WO200077040-A2

PD 21-DEC-2000.

PF 16-JUN-2000; 2000WO-US166636.

PR 16-JUN-1999; 99US-0139566.

PR 09-NOV-1999; 99US-0164417.

PA (INCY-) INCYTE GENOMICS INC.

PI Yue H, Tang YT, Hillman JL

XX 0001-035374-003

DR N-PSDB; AAF32646.

PT New human intrace

PT inflammatory, neurological, gastrointestinal, reproductive and developmental disorders -

PS Claim 5; Page 120-121; 192pp; English

CC Sequences AAF32638 - AAF32659 represent cDNA encoding human
CC intracellular signalling molecules INTRA1 - INTRA52, represented in
CC AAB64369 - AAB64420. Modulators of the intracellular signalling molecules
CC of the invention exhibit immunosuppressive; cytostatic; neuroprotective;
CC neurotropic; antiarteriosclerotic; antiinflammatory; anti-HIV;
CC neuroleptic; antibacterial; antifungal; antiviral; polypeptidic;
CC antihelminthic; and antiparasitismal activity. INTRA polypeptides their
CC agonists and antagonists are useful for the treatment of a condition
CC associated with decreased or increased expression of functional INTRA,
CC disorders associated with abnormal INTRA expression or activity include
CC cell, proliferative disorders e.g. arteriosclerosis and cancers;
CC auto-immune or inflammatory disorders e.g. Addison's disease and acquired
CC immunodeficiency syndrome (AIDS); viral, bacterial, fungal, parasitic,
CC protozoal and helminthic infections; gastrointestinal disorders e.g.
CC dysphagia and irritable bowel syndrome; neurological disorders e.g.
CC epilepsy and Parkinson's disease; piton diseases e.g. Creutzfeldt-Jakob
CC disease and mental disorders e.g. anxiety, schizophrenia and Tourette's
CC disorder. Antibodies immuno specific for the INTRA proteins may also be
CC useful in the diagnosis of the above disorders.

5Q Sequence 319 AA;

Query Match	41.08;	Score 679.5;	DB 22;	Length 319;
PostgreSQL 9.5.12	42.38;	Score 679.5;	DB 22;	Length 319;

Matches 142; Conservative 48; Mismatches 87; Indels 51; Gaps 9;

```

0Y      62  TCRKTRGGAPRLA-----TPARCDODVTHHHSF-----92
      11  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      55  egsrvvckrepactpqrgekherethellprlrvnnnyghrhtslghnavlpsnargllsrs 114
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
0Y      93  --SGRRSSVSSSSSSSSDRLDHHMAP--PVADQASPAKRVTPQKRVNCQPLDLKGPVP 149
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

```

Dd	115	tstgsaaagsgnsaaasegglllgisrpttprpyphrhserairtqrqjllvd--dlkg-slk	172
Oy	150	PELDKNHLLDEAGCKKCKCCKACAPRFLPSCWCNCNDCCLSAOTLVNKGTCMKVLVGIRYH	209
Dd	172	edlqfghflceegcgckcckgeactaprflpsclacnrgqlcsaesmweygtcmclvkglfyh	231
Oy	210	CTNFDDGSGCADHPHCSCSRNSCNCAFMSPFGALSVLPCLCYLPTMGCWKLARCGYDRLR	269
Dd	232	csmddgedgdsydnpccscsgshccstyrlcmgamsilfpcllycpbakgklkicrtecydwih	291
Oy	270	RPGCRKHNTSNVICKAAS---GDAKTS	293
Dd	292	rpgcrcksnntvycklescpsrqgdkps	319
RESULT	3		
AAM40474	ID	AAM40474 standard; Protein: 326 AA.	
XX			
AC		AAM40474;	
Dt		22-OCT-2001 (first entry)	
XX			
DE		Human polypeptide SEQ ID NO 5405.	
XX			
KW		Human; nootropic; immunosuppressant; cyostatic; gene therapy; cancer;	
KW		peripheral nervous system; neuropathy; central nervous system; CNS;	
KW		Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KW		amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
KW		chemokineic; thrombolytic; drug screening; arthritis; inflammation;	
KW		leukaemia.	
OS		Homo sapiens.	
XX			
PN		WO200153312-A1.	
XX			
PD		26-JUL-2001.	
XX			
PF		26-DEC-2000; 2000MO-US34263.	
XX			
PR		21-JAN-2000; 2000US-0488725.	
PR		25-APR-2000; 2000US-0552317.	
PR		09-JUL-2000; 2000US-0598042.	
PR		19-JUL-2000; 2000US-0620312.	
PR		03-AUG-2000; 2000US-0653450.	
PR		14-SEP-2000; 2000US-0662191.	
PR		19-OCT-2000; 2000US-0693036.	
PR		29-NOV-2000; 2000US-0727344.	
XX			
PA		(HYSE-) HYSEQ INC.	
PI		Tang Y.T., Liu C., Asundi V., Chen R., Ma Y., Qian X.B., Ren F., Wang D.	
PI		Wang J., Wang Z., Wehrman T., Xu C., Xue A.J., Qian Y., Zhang J.	
PI		Zhao Q.A., Zhou P., Goodrich R., Drmannic RT;	
XX			
DR		WPI: 2001-442253/47.	
N-PSDB:		AAI59630.	
PT		Novel nucleic acids and polypeptides, useful for treating disorders	
XX		such as central nervous system injuries -	
XX			
PS		Example 2; SEQ ID NO 5405; 10078bp; English.	
XX			
CC		The invention relates to human nucleic acids (AAI57798-AAI61369) and	
CC		the encoded polypeptides (AAM38642-AAM42213) with nootropic,	
CC		immunosuppressant and cyostatic activity. The polynucleotides are used	
CC		in gene therapy. A composition containing a polypeptide or polynucleo-	
CC		of the invention may be used to treat diseases of the peripheral nerve	
CC		system, such as peripheral nervous injuries, peripheral neuropathy and	
CC		localised neuropathies and central nervous system diseases, such as	
CC		Alzheimer's, Parkinson's disease, Huntington's disease, amyotropic	
CC		lateral sclerosis, and Shy-Drager Syndrome. Other uses include the	
CC		utilisation of the activities such as: Immune system suppression,	

	CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
	CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
	CC	assays for receptor activity, arthritis and inflammation, leukaemia and
	CC	C.N.S disorders.
	CC	Note: The sequence data for this patent did not form part of the printed
	CC	specification.
	XX	
SQ	Sequence	326 AA;
	Query Match	41.0%; Score 679.5; DB 22; Length 326;
	Best Local Similarity	43.3%; Pred. No. 6..le-49;
	Matches 142;	Conservative 48; Mismatches 87; Indels 51; Gaps 9
OY	6 POSADLPNNSWV--OPLIDS--RMSHSLQHPLTLPIDQKTSHEVDNDPSLATL 61	
DB	10 pndqsgssglvldgprlsdrpidyrelgdptalsldqikatrsgney-----t 61	
OY	62 TPKPKTRGCAPELAR-----TPARCDQDVTHMHSF----- 92	
DB	62 egrpsvvkkrpaptarpeqekherthelipinvnmvehrtshlghavlrsnargpllrts 121	
OY	93 --SGRRSSVSSTSSSTDORLDLHMHP--PVADOASPAVRIOQKVHOCPLDLKPAPV 149	
DB	122 tctgsaaasgsmasasagllgrsprrtpvrphrgseratrlrpkqlvd--dlkg-slk 178	
OY	150 PELDKHFLLCEACGKCCKECASPRTLPGCWVCNQDCLCSAOTLVNYGTMCCLVGIFTH 209	
DB	179 edltqltkficedqgkykcgcetctarlfpclaclcnrcjclcsaesmvegytcmlvk1fyh 238	
OY	210 CPMNEDEGCADHPGSCSNCCARBFPGALSVLPCLLCYVPATGCVKLAORGVDRLR 269	
DB	239 csnddegdsydmdpcscsqshccstylnmgamslfipcllcypbakvgclklcrtygwln 298	
OY	270 RFGCRCKHTNSTYICKAAS---GDAKTS 293	
DB	299 rpgcrckmsntvycklescpsrqgkps 326	
RESULT 4		
AAV87591		
ID AAV87591	standard; Protein: 319 AA.	
XX AC AAV87591:		
XX DE 21-JUL-2000	(first entry)	
XX DE Human SPROUTY-1 protein, SEQ ID NO:24.		
XX KW SPROUTY-1; spryl; fibroblast growth factor; FGF inhibitor;		
KW signalling; tumorigenesis; neurogenesis; metanephrogenesis; cancer;		
KW tumour; prostate; adrenal gland.		
XX OS Homo sapiens.		
XX FH Key Location/Qualifiers		
FT Misc-difference 173..239		
FT FT Note="The base sequence encoding these residues is		
FT FT absent from this protein's corresponding cDNA		
FT FT sequence"		
PN WO200015781-A1.		
PD 23-MAR-2000.		
PX 10-SEP-1999;	99WO-US20638.	
PR 11-SEP-1998;	98US-009936.	
PA (ELIL) LILLY & CO ELI.		
PI Edmonds BT;		
XX		

DR WPI: 2000-271422/23.
DR N-PSDB: AAA14071.

PT New polynucleotide encoding a 319 amino acid sequence, used as a
PT fibroblast growth factor (FGF) antagonist for treating prostate cancer
PT and inhibiting angiogenesis in mammals -

XX Claim 1; Fig 26; 93pp; English.

XX This sequence represents a human SPROUTY-1 protein (also referred
XX to as spry1 in the specification). SPROUTY-1 proteins (AA187591-187595)
XX are antagonists of members of the fibroblast growth factor (FGF)
XX family. The SPROUTY-1 proteins of the invention, and fragments
XX thereof may be used in mammals to limit FGF levels in vivo and
XX to control FGF-mediated signaling. They may be used to counteract FGF
XX overexpression associated with certain types of tumors, and to control
XX the growth, development or differentiation of any cell responsive to
XX FGF-mediated signaling. For example, SPROUTY-1 proteins may be used to
XX control tumorigenesis, neurogenesis, metanephrogenesis, and particularly
XX SPROUTY-1 underexpression including polycystic kidney disease,
XX nephropathy, lung fibrosis, hepatic fibrogenesis, endometrial
XX insufficiency and various cancers, particularly prostatic tumours.
XX They may additionally be used to treat a mammal with an adrenal gland
XX pathology involving hypersensitive responses to FGF. Nucleic acid
XX sequences encoding SPROUTY-1 proteins may be used to recombinantly
XX express SPROUTY-1 proteins, as a source of primers and probes, and for
XX antisense suppression of SPROUTY-1 gene expression.

SO Sequence 319 AA;

Query Match 40.6%; Score 672.5; DB 21; Length 319;
Best Local Similarity 43.0%; Pred. No. 2.3e-48;
Matches 141; Conservative 47; Mismatches 89; Indels 51; Gaps 9;

OY 6 POSAPLPNSVMV--OPLDLS--RMSHSRLQHPITLPIQOVKTSVNDYIDNPSIALT 61
DB 3 pndqngsgsslvvlgqpsldsrqldyereipqalalsldqikayrsney-----t 54
OY 62 TGPRTRGAPBELAP-----TPARCDDVTTHMISF----- 92
DB 55 egpavvkrpaprtaqrqekhertheliprvnmnyehrhshghavlpnargplslrs 114
OY 93 ---SGRPSSVSSSTSSDRLDLMHAPP--PVADQASPAVRIRQPKVHCQPLDLKGPAPV 149
DB 115 tsfgsaassgssnsaseqgllgrspptprpghrselaltqpkqlvd--dlkg-slk 171
OY 150 PELDKHFLLEACGCKCKECASPRTLPSGCWVNOECLSAOTLVNNGTMCMLVQGIFFH 209
DB 172 editqhkfcceqgckckgectaprlpsclacnrgclsaesmveygtcmclvpglflfh 231
OY 210 CTNEDDEGSCADHPDCSRSRNCARMSFGALSVLPLCLLYLPATGCVKLAQGRYDLRL 269
DB 232 csnddegdsyadnpscsqshncserylcmgamsflpcollycpakacklrrrcywh 291
OY 270 RPRGRCKHTNSVTRKAS-----GDAKTS 293
DB 292 rpgcrcknntlycklscspstgqgkps 319

RESULT 5

AA187594 standard: Protein; 338 AA.

AA187594;

XX 21-JUL-2000 (first entry)

DE Human SPROUTY-1 protein, SEQ ID NO:7.

XX SPROUTY-1; spry1; fibroblast growth factor; FGF inhibitor;
XX signalling; tumorigenesis; neurogenesis; metanephrogenesis; cancer;

KW tumour; prostate; adrenal gland.

OS Homo sapiens.

XX Key Location/Qualifiers

XX MISC-difference 191..258

FT /note="The base sequence encoding these residues is
FT absent from this protein's corresponding cDNA
FT sequence"

XX W0200015781-A1.

XX 23-MAR-2000.

XX 10-SEP-1999; 99MO-US20638.

XX 11-SEP-1998; 98US-0099936.

XX (BLIL) LILLY & CO ELI.

XX Edmonds BT;

XX WPI: 2000-271422/23.

XX N-PSDB: AAA14074.

PT New polynucleotide encoding a 319 amino acid sequence, used as a
PT fibroblast growth factor (FGF) antagonist for treating prostate cancer
PT and inhibiting angiogenesis in mammals -

XX Example 4; Fig 7; 93pp; English.

XX This sequence represents a human SPROUTY-1 protein (also referred
XX to as spry1 in the specification). SPROUTY-1 proteins (AA187591-187595)
XX are antagonists of members of the fibroblast growth factor (FGF)
XX family. The SPROUTY-1 proteins of the invention, and fragments
XX thereof may be used in mammals to limit FGF levels in vivo and
XX to control FGF-mediated signaling. They may be used to counteract FGF
XX overexpression associated with certain types of tumors, and to control
XX the growth, development or differentiation of any cell responsive to
XX FGF-mediated signaling. For example, SPROUTY-1 proteins may be used to
XX control tumorigenesis, neurogenesis, metanephrogenesis, and particularly
XX SPROUTY-1 underexpression including polycystic kidney disease,
XX nephropathy, lung fibrosis, hepatic fibrogenesis, endometrial
XX insufficiency and various cancers, particularly prostatic tumours.
XX They may additionally be used to treat a mammal with an adrenal gland
XX pathology involving hypersensitive responses to FGF. Nucleic acid
XX sequences encoding SPROUTY-1 proteins may be used to recombinantly
XX express SPROUTY-1 proteins, as a source of primers and probes, and for
XX antisense suppression of SPROUTY-1 gene expression.

SO Sequence 338 AA;

Query Match 40.3%; Score 668.5; DB 21; Length 338;
Best Local Similarity 43.0%; Pred. No. 5.3e-48;
Matches 141; Conservative 48; Mismatches 88; Indels 51; Gaps 9;

OY 6 POSAPLPNSVMV--OPLDLS--RMSHSRLQHPITLPIQOVKTSVNDYIDNPSIALT 61
DB 22 pndqngsgsslvvlgqpsldsrqldyereipqalalsldqikayrsney-----t 73
OY 62 TGPRTRGAPBELAP-----TPARCDDVTTHMISF----- 92
DB 74 egpavvkrpaprtaqrqekhertheliprvnmnyehrhshghavlpnargplslrs 133
OY 93 ---SGRPSSVSSSTSSDRLDLMHAPP--PVADQASPAVRIRQPKVHCQPLDLKGPAPV 149
DB 134 tsfgsaassgssnsaseqgllgrspptprpghrselaltqpkqlvd--dlkg-slk 190
OY 150 PELDKHFLLEACGCKCKECASPRTLPSGCWVNOECLSAOTLVNNGTMCMLVQGIFFH 209
DB 191 editqhkfcceqgckckgectaprlpsclacnrgclsaesmveygtcmclvpglflfh 250

QY 210 CTNEDBGCADHPCSCSNCARMSFMGALSVLPCLLCYTPATGCVKLAORGVDRLR 269
 DB 251 csndegdsysdnpccscgchccsrylcmgamsllfpcllcyppakgclklctrcydwln 310
 QY 270 RPGCRCKHTNSVICRKAAS----GDAKTS 293
 DB 311 rfgrcknntvcklscscpsrgqkps 338

RESULT 6
 AAY87592
 ID AAY87592 standard; Protein; 294 AA.
 AC AAY87592;
 DT 21-JUL-2000 (first entry)
 DE Human SPROUTY-1 protein, SEQ ID NO:3.
 XX
 KW SPROUTY-1; spry1; fibroblast growth factor; FGF inhibitor;
 KW signaling; tumorigenesis; neurogenesis; metanephrogenesis; cancer;
 KW tumour; prostate; adrenal gland.
 XX
 OS Homo sapiens.
 PN WO200015781-A1.
 PD 23-MAR-2000.
 XX
 PF 10-SEP-1999; 99WO-US20638.
 XX
 PR 11-SEP-1998; 98US-0099936.
 XX
 PA (ELIL) LILLY & CO ELI.
 XX
 PI Edmonds BT;
 DR WPI: 2000-271422/23.
 DR N-PSDB; AAA14072.
 XX
 PT New polynucleotide encoding a 319 amino acid sequence, used as a
 PT fibroblast growth factor (FGF) antagonist for treating prostate cancer
 PT and inhibiting angiogenesis in mammals -
 XX
 PS Disclosure: Fig 3; 93pp; English.
 XX
 CC This sequence represents a human SPROUTY-1 protein (also referred
 CC to as spry1 in the specification). SPROUTY-1 proteins (AAY87591-187595)
 CC are antagonists of members of the fibroblast growth factor (FGF)
 CC family. The SPROUTY-1 proteins of the invention, and fragments
 CC thereof may be used in mammals to limit FGF levels in vivo and
 CC to control FGF-mediated signaling. They may be used to counteract FGF
 CC overexpression associated with certain types of tumours, and to control
 CC the growth, development or differentiation of any cell responsive to
 CC FGF-mediated signalling. For example, SPROUTY-1 proteins may be used to
 CC control tumorigenesis, neurogenesis, metanephrogenesis, and particularly
 CC angiogenesis. They may also be used to treat conditions associated with
 CC SPROUTY-1 underexpression including polycystic kidney disease,
 CC nephropathy, lung fibrosis, hepatic fibrogenesis, endometrial
 CC insufficiency and various cancers, particularly prostatic tumours.
 CC They may additionally be used to treat a mammal with an adrenal gland
 CC pathology involving hypersensitive responses to FGF. Nucleic acid
 CC sequences encoding SPROUTY-1 proteins may be used to recombinantly
 CC express SPROUTY-1 proteins, as a source of primers and probes, and for
 CC antisense suppression of SPROUTY-1 gene expression.
 XX
 SQ Sequence 294 AA;

Query Match 39.5%; Score 654; DB 21; Length 294;
 Best Local Similarity 44.2%; Pred. No. 7, 3e-47;
 Matches 133; Conservative 42; Mismatches 78; Indels 48; Gaps 8;

QY 30 RLQ-HEPLTLPIDQVKTSHVENDYIDNBSLALTGPCKRTGCAPELAP-----T 77
 DB 5 rlelqbtallslsdqkalkatrsney-----tegpsvwxkrpaptaprekeherthei 56
 QY 78 PARCDDVTYHHMISF-----SGRRSSVSSSSSSSDORLIDHMAP 117
 DB 57 pluvnnnyehrlhtshlghavlpsnargpllsrstcgsaassgsrnsseqgllgrspp 116
 QY 118 P-PVADQASPRAVRIQPKVYHCOPDLKGPVAPPELDKHFLLCEACGKCKEACSPRTL 176
 DB 117 trpyghrseratrlrqpklvld--dlkg-slkedltqkflficeqgckcgcaccaptl 173
 QY 177 PSCWVCNDEPLCSAGPTLVNFGCMCLVQGIHYHCTNEDDEGSCADHPCSCSNCARMS 236
 DB 174 psclacnrgqlcsaeswvegtcmclvkgifhcsndegdsysdnpccscgchccsryl 233
 QY 237 FMGALSVLPCLLCYTPATGCVKLAORGVDRLRPGCRCKHTNSVICRKAAS----GDAKT 292
 DB 234 cmgamsllfpcllcyppakgclklctrcydwlnrpgrcknntvcklscscpsrgqkps 293
 QY 293 S 293
 DB 294 s 294

RESULT 7
 AAY87597
 ID AAY87597 standard; Protein; 288 AA.
 AC AAY87597;
 DT 21-JUL-2000 (first entry)
 DE Human SPROUTY protein long consensus sequence.
 XX
 KW SPROUTY; spry; fibroblast growth factor; FGF inhibitor;
 KW signaling; tumorigenesis; neurogenesis; metanephrogenesis; cancer;
 KW tumour; prostate; adrenal gland; consensus sequence.
 XX
 OS Homo sapiens.
 PN WO200015781-A1.
 PD 23-MAR-2000.
 XX
 PF 10-SEP-1999; 99WO-US20638.
 XX
 PR 11-SEP-1998; 98US-0099936.
 XX
 PA (ELIL) LILLY & CO ELI.
 XX
 PI Edmonds BT;
 DR WPI: 2000-271422/23.
 XX
 PT New polynucleotide encoding a 319 amino acid sequence, used as a
 PT fibroblast growth factor (FGF) antagonist for treating prostate cancer
 PT and inhibiting angiogenesis in mammals -
 XX
 PS Disclosure: Fig 2; 93pp; English.
 XX
 CC Sequences AAY87596-V87597 represent, respectively, short and long
 CC SPROUTY protein consensus sequences. The invention relates to human
 CC SPROUTY-1 proteins (also referred to as spry1 in the specification);
 CC AAY87591-AAY87595) which are antagonists of members of the fibroblast
 CC growth factor (FGF) family. The SPROUTY-1 proteins of the invention, and
 CC fragments thereof may be used in mammals to limit FGF levels in vivo and
 CC to control FGF-mediated signalling. They may be used to counteract FGF
 CC overexpression associated with certain types of tumours, and to control
 CC the growth, development or differentiation of any cell responsive to
 CC FGF-mediated signalling. For example, SPROUTY-1 proteins may be used to
 CC control tumorigenesis, neurogenesis, metanephrogenesis, and particularly

CC angiogenesis. They may also be used to treat conditions associated with
 CC SPROUTY-1 underexpression including polycystic kidney disease,
 CC nephropathy, lung fibrosis, hepatic fibrogenesis, endometrial
 CC insufficiency and various cancers, particularly prostatic tumours.
 CC They may additionally be used to treat a mammal with an adrenal gland
 CC pathology involving hypersensitive responses to FGF. Nucleic acid
 CC sequences encoding SPROUTY-1 proteins may be used to recombinantly
 CC express SPROUTY-1 proteins, as a source of primers and probes, and for
 CC antisense suppression of SPROUTY-1 gene expression.

XX Sequence 288 AA;

Query Match 39.4%; Score 653.5; DB 21; Length 288;
 Best Local Similarity 44.3%; Pred. No. 7.8e-47;
 Matches 131; Conservative 41; Mismatches 77; Indels 47; Gaps 7;

QY 34 PLTLPIDQVKTSHENDYIDNPSLALTTGPKRTGCAPELAP-----TPARCD 82
 DB 4 PLTLPIDQVKTSHENDYIDNPSLALTTGPKRTGCAPELAP-----TPARCD 82
 QY 83 QDVTHHMTSF-----SGRPSSVSSSTSSDORLIDHMAP-PVA 121
 DB 56 nmyehrheshghavlpnarnatgpllrstscgsaasgsnsasasegllgrspptprvp 115
 QY 122 DQASRAVRIQPKVHCQPLDLKGPVAPPELDKHFLLCEACGCKCKECASPRTLPCWV 181
 DB 116 ghrserairtrpqlvnd-dlkg-slkedltqhkflcegcgckgcgecactrlpscla 172
 QY 182 CNOECILCSAOTLVNNGCMCLVGGIFVHCTNEDDEGSCADHPGCSRSNCCARMSFGAL 241
 DB 173 cnrgclcsaesmveglcmclvkgllfychnddegdsydnpcscsqhccsrylcmgam 232
 QY 242 SVVLPCLCYLPATGCVKLAQRGYDRLRRPGCRCKHNSVTCGAAS---GDAKTS 293
 DB 223 slflpcllcyppakgclklcrctcydwlhnpgrcknntvcklescpstrgqkps 288

RESULT 8
 AAY87593
 ID AAY87593 standard; Protein: 295 AA.

XX AAY87593;

DT 21-JUL-2000 (first entry)

DE Human SPROUTY-1 protein, SEQ ID NO:5.

XX SPROUTY-1; spry1; fibroblast growth factor; FGF inhibitor;
 KW signalling; tumorigenesis; neurogenesis; metanephrogenesis; cancer;
 KM tumour; prostate; adrenal gland.

XX Homo sapiens.

OS WO200015781-A1.

PN 23-MAR-2000.

PF 10-SEP-1999; 99WO-US20638.

PR 11-SEP-1998; 98US-0099936.

PA (ELIL) LILLY & CO ELI.

PI Edmonds BT;

XX WPI: 2000-271422/23.

DR N-PSDB; AAA14073.

XX New polynucleotide encoding a 319 amino acid sequence, used as a
 PT fibroblast growth factor (FGF) antagonist for treating prostate cancer
 PT and inhibiting angiogenesis in mammals -
 XX

PS Example 4; Fig 5; 93pp; English.

XX This sequence represents a human SPROUTY-1 protein (also referred
 CC to as spry1 in the specification). SPROUTY-1 proteins (AAY87591-787995)
 CC are antagonists of members of the fibroblast growth factor (FGF)
 CC family. The SPROUTY-1 proteins of the invention, and fragments
 CC thereof may be used in mammals to limit FGF levels in vivo and
 CC to control FGF-mediated signalling. They may be used to counteract FGF
 CC overexpression associated with certain types of tumours, and to control
 CC the growth, development or differentiation of any cell responsive to
 CC FGF-mediated signalling. For example, SPROUTY-1 proteins may be used to
 CC control tumorigenesis, neurogenesis, metanephrogenesis, and particularly
 CC angiogenesis. They may also be used to treat conditions associated with
 CC nephropathy, lung fibrosis, hepatic fibrogenesis, endometrial
 CC insufficiency and various cancers, particularly prostatic tumours.
 CC They may additionally be used to treat a mammal with an adrenal gland
 CC pathology involving hypersensitive responses to FGF. Nucleic acid
 CC sequences encoding SPROUTY-1 proteins may be used to recombinantly
 CC express SPROUTY-1 proteins, as a source of primers and probes, and for
 CC antisense suppression of SPROUTY-1 gene expression.

XX Sequence 295 AA;

Query Match 39.4%; Score 653.5; DB 21; Length 295;
 Best Local Similarity 44.3%; Pred. No. 8.1e-47;
 Matches 131; Conservative 41; Mismatches 77; Indels 47; Gaps 7;

QY 34 PLTLPIDQVKTSHENDYIDNPSLALTTGPKRTGCAPELAP-----TPARCD 82
 DB 11 PLTLPIDQVKTSHENDYIDNPSLALTTGPKRTGCAPELAP-----TPARCD 82
 QY 83 QDVTHHMTSF-----SGRPSSVSSSTSSDORLIDHMAP-PVA 121
 DB 63 nmyehrheshghavlpnarnatgpllrstscgsaasgsnsasasegllgrspptprvp 122
 QY 122 DQASRAVRIQPKVHCQPLDLKGPVAPPELDKHFLLCEACGCKCKECASPRTLPCWV 181
 DB 123 ghrserairtrpqlvnd-dlkg-slkedltqhkflcegcgckgcgecactrlpscla 179
 QY 182 CNOECILCSAOTLVNNGCMCLVGGIFVHCTNEDDEGSCADHPGCSRSNCCARMSFGAL 241
 DB 180 cnrgclcsaesmveglcmclvkgllfychnddegdsydnpcscsqhccsrylcmgam 239
 QY 242 SVVLPCLCYLPATGCVKLAQRGYDRLRRPGCRCKHNSVTCGAAS---GDAKTS 293
 DB 240 slflpcllcyppakgclklcrctcydwlhnpgrcknntvcklescpstrgqkps 295

RESULT 9
 AAY8792
 ID AAY8792 standard; Protein: 315 AA.

XX AAY8792;

DT 11-SEP-1998 (first entry)

DE Homo sapiens sprouty 2 protein.

XX spry2; sprouty 2 protein; angiogenesis; treatment; cancer;
 KW lateral inhibitory signal; epithelial tube branching; inhibition;
 KW neovascularisation; diabetic retinopathy; retrorenal fibroplasia;
 KW neovascular glaucoma; psoriasis; angiodermas; immune; non-immune;
 KW inflammation; rheumatoid arthritis; capillary proliferation;
 KW atherosclerotic plaques; haemangiomas; Kaposi's Sarcoma;
 KW endometriosis; scar formation; wound healing.

XX Homo sapiens.

PN WO9820032-A1.

XX 14-MAY-1998.

```

XX 07-NOV-1997; 97WO-US20398.
XX
XX 07-NOV-1996; 96US-0030232.
XX
XX (STRD) UNIV LELAND STANFORD JUNIOR.
XX
XX Hachem N, Krasnow MA;
XX WPI; 1998-286862/25.
XX DR N-PSDB; AAV32399.
XX
XX New isolated human sprouty polypeptide(s) - used to develop products
XX for treating disorders involving angiogenic dysfunction, e.g.
XX cancer, diabetic retinopathy, glaucoma, psoriasis or inflammation
XX
XX Claim 3; Page 55-56; 90pp; English.
XX
XX The sequence is that of human sprouty 2 protein. The protein
XX functions as a lateral inhibitory signal during epithelial tube
XX branching. It can be used for inhibiting angiogenesis and
XX neovascularisation in the treatment of e.g. cancers, diabetic
XX angiopathy, retrolental fibroplasia, neovascular glaucoma, psoriasis,
XX angiofibrosis, immune and non-immune inflammation (including rheumatoid
XX arthritis), capillary proliferation within atherosclerotic plaques,
XX haemangiomas, Kaposi's Sarcoma, endometriosis and unwanted scar
XX formation in wound healing. Products of the DNA and proteins
XX can also be used for detection and diagnosis.
XX
XX Sequence 315 AA;
XX
XX Query Match 38.7%; Score 641.5; DB 19; Length 315;
XX Best Local Similarity 43.6%; Pred. No. 8.7e-46;
XX Matches 130; Conservative 45; Mismatches 84; Indels 39; Gaps 8;
XX
XX 19 QPILDSRMHSR-----LQHPLTLPRIDQVTSHEVDYIDNPSLALTGK--- 65
XX 12 qplldgrldggrgqgpdprdalqgvnlslqdlralntneytegrvdrpqlkpp 71
XX
XX 66 -----RTRCGAPELATPP-----ARCDQVTHNMFSGRPSVSSSSSTSS 107
XX 72 rpsdqkherlglpdehrrpdlqhsqvsaraplsistvssgstststsssss 131
XX
XX 108 DORLL-DHMAAPPVADQASPRAVRIQPKVHCCQPLDLKPAVPELDKHFLLCEACGKCK 166
XX 132 eqlllgssfsqpvad---gltrvgpk-selkpgelk-plakediglhaycedecgkck 185
XX
XX 167 CRECAPRRLPSQWVNOECISAOFLVNYGTCMCLVOCIFHCTNEDEGSCADHPGSC 226
XX 186 ckectyprrlpdpwldckqclcsagvldygtccvkgllfychsn-dedcncadnpcc 244
XX
XX 227 SRSCCARSPFMGALSVLPCLCLCYLPATGCVKLAORGYDRLRKPRGCRCKHTNSVICK 284
XX 245 sqshoctlrswangymalfplclwcyipakgclklcgqcydrvmrpgcrcknntvck 302
XX
XX
XX RESULT 10
XX AA87595
XX ID AAY87595 standard; Protein: 167 AA.
XX
XX AC AAY87595;
XX
XX 21-JUL-2000 (first entry)
XX
XX Human SPROUTY-1 protein, SEQ ID NO:9.
XX
XX DE Human SPROUTY-1; spyr1; fibroblast growth factor; FGF inhibitor;
XX KM signalling; tumorigenesis; neurogenesis; metanephrogenesis; cancer;
XX KW tumour; prostate; adrenal gland.
XX
XX OS Homo sapiens.
XX

```

```

XX PN WO200015781-A1.
XX PD 23-MAR-2000.
XX
XX PF 10-SEP-1999; 99WO-US20638.
XX
XX PR 11-SEP-1998; 98US-0099936.
XX
XX PA (ELIL) LILLY & CO ELI.
XX
XX PI Edmonds BT;
XX
XX DR WPI; 2000-271422/23.
XX DR N-PSDB; AAA14075.
XX
XX New polynucleotide encoding a 319 amino acid sequence, used as a
XX fibroblast growth factor (FGF) antagonist for treating prostate cancer
XX and inhibiting angiogenesis in mammals -
XX
XX Example 4; Fig 9; 93pp; English.
XX
XX This sequence represents a human SPROUTY-1 protein (also referred
XX to as spyr1 in the specification). SPROUTY-1 proteins (AA87591-Y87595)
XX are antagonists of members of the fibroblast growth factor (FGF)
XX family. The SPROUTY-1 proteins of the invention, and fragments
XX thereof may be used in mammals to limit FGF levels in vivo and
XX to control FGF-mediated signalling. They may be used to counteract FGF
XX overexpression associated with certain types of tumours, and to control
XX the growth, development or differentiation of any cell responsive to
XX FGF-mediated signalling. For example, SPROUTY-1 proteins may be used to
XX control tumorigenesis, neurogenesis, metanephrogenesis, and particularly
XX angiogenesis. They may also be used to treat conditions associated with
XX SPROUTY-1 underexpression including polycystic kidney disease,
XX nephropathy, lung fibrosis, hepatic fibrogenesis, endometrial
XX insufficiency and various cancers, particularly prostatic tumours.
XX They may additionally be used to treat a mammal with an adrenal gland
XX pathology involving hypersensitive responses to FGF. Nucleic acid
XX sequences encoding SPROUTY-1 proteins may be used to recombinantly
XX express SPROUTY-1 proteins, as a source of primers and probes, and for
XX antisense suppression of SPROUTY-1 gene expression.
XX
XX Sequence 167 AA;
XX
XX Query Match 33.6%; Score 557.5; DB 21; Length 167;
XX Best Local Similarity 57.1%; Pred. No. 4.3e-39;
XX Matches 96; Conservative 32; Mismatches 31; Indels 9; Gaps 3;
XX
XX 130 RIQRPVNHQRPDLKGPVAPPELDKHFLLCEACGKCKEASPTLPSQWVNOECICS 189
XX 5 rvrprvr-----dlkg-slkedltqhkfticcgqckcgcetaprltspclacmqclcs 59
XX
XX 190 AQLVNYGTCMCLVOCIFHCTNEDEGSCADHPGSCSRSCNCCARMSFGALSVLPCLL 249
XX 60 aeqmweygtcmclvgyflfychsncddegdsydnpscsqshoccsyltmgamsltlpol 119
XX
XX 250 CYLPATGCVKLAORGYDRLRKPRGCRCKHTNSVICAAS---GDAKTS 293
XX 120 cypakgclklorcydwthrpgrccknsntvcklscscpryggkps 167
XX
XX
XX RESULT 11
XX AA87596
XX ID AAY87596 standard; Protein: 156 AA.
XX
XX AC AAY87596;
XX
XX 21-JUL-2000 (first entry)
XX
XX Human SPROUTY protein short consensus sequence.
XX
XX DE Human SPROUTY; spyr; fibroblast growth factor; FGF inhibitor;
XX KM signalling; tumorigenesis; neurogenesis; metanephrogenesis; cancer;
XX

```

KW tumour; prostate; adrenal gland; consensus sequence.
 XX
 OS Homo sapiens.
 XX
 PN WO200015781-A1.
 XX
 PD 23-MAR-2000.
 XX
 PF 10-SEP-1999; 99WO-US20638.
 XX
 PR 11-SEP-1998; 98US-0099936.
 XX
 PA (ELIL) LILLY & CO ELI.
 XX
 PI Edmonds BT;
 XX
 DR WPI: 2000-271422/23.
 XX
 PT New polynucleotide encoding a 319 amino acid sequence, used as a
 PT fibroblast growth factor (FGF) antagonist for treating prostate cancer
 PT and inhibiting angiogenesis in mammals -
 XX
 PS Disclosure; Fig 1; 93pp; English.

XX
 CC Sequences AAY87596-Y87597 represent, respectively, short and long
 CC SPROUTY protein consensus sequences. The invention relates to human
 CC SPROUTY-1 proteins (also referred to as spry1 in the specification;
 CC AAY87591-AAY87595) which are antagonists of members of the fibroblast
 CC growth factor (FGF) family. The SPROUTY-1 proteins of the invention, and
 CC fragments thereof may be used in mammals to limit FGF levels in vivo and
 CC to control FGF-mediated signalling. They may be used to counteract FGF
 CC overexpression associated with certain types of tumours, and to control
 CC the growth, development or differentiation of any cell responsive to
 CC FGF-mediated signalling. For example, SPROUTY-1 proteins may be used to
 CC control tumorigenesis, neurogenesis, metanephrogenesis, and particularly
 CC angiogenesis. They may also be used to treat conditions associated with
 CC SPROUTY-1 underexpression including polycystic kidney disease,
 CC nephropathy, lung fibrosis, hepatic fibrogenesis, endometrial
 CC insufficiency and various cancers, particularly prostatic tumours.
 CC They may additionally be used to treat a mammal with an adrenal gland
 CC pathology involving hypersensitive responses to FGF. Nucleic acid
 CC sequences encoding SPROUTY-1 proteins may be used to recombinantly
 CC express SPROUTY-1 proteins, as a source of primers and probes, and for
 CC antisense suppression of SPROUTY-1 gene expression.
 CC
 XX
 SQ Sequence 156 AA;

Query Match 32.8%; Score 543.5; DB 21; Length 156;
 Best Local Similarity 59.0%; Pred. No. 5.9e-38;
 Matches 92; Conservative 29; Mismatches 30; Indels 5; Gaps 2;
 QY 142 DLKSPAVPEPLDKHFLCEACGCKCKEASPTLPSWCWVNCQECISADTLVNYGCMC 201
 Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
 QY 202 LVGGIFVHCTNEDDEGSCADHPSCSRNSNCARNSFGALSVVLPCLICTYLPATGCVKLA 261
 Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
 QY 61 lvygifyfhcnddegdsydnpcscsqshcsrylcmgamsilflpollycpypakgcklkic 120
 Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
 QY 262 ORGDRRLRRGCRCKRHNSVYCKAAS---GDAKTS 293
 Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
 Db 121 rfcydwihpgrctcknsntvycklescpstvggkps 156

RESULT 12
 AAB41184
 ID AAB41184 standard; Protein; 139 AA.
 AC AAB41184;
 XX
 DT 08-FEB-2001 (first entry)
 XX

DE Human ORFX ORF948 polypeptide sequence SEQ ID NO:1896.
 XX
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnerable; antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotrophic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antihemetic; antihypertensive;
 KW antinaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 XX
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 XX
 PR 02-APR-1999; 99US-0127636.
 XX
 PR 05-APR-1999; 99US-0127728.
 XX
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR WPI: 2000-602362/57.
 XX
 DR N-PDB: AAC75393.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11; Page 1453; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnerable;
 CC antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotrophic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antihemetic;
 CC antihypertensive and antinaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 139 AA;

Query Match 31.9%; Score 528; DB 21; Length 139;
 Best Local Similarity 62.6%; Pred. No. 1e-36;
 Matches 87; Conservative 24; Mismatches 24; Indels 4; Gaps 1;
 QY 159 CEACGCKCKEACSPRTLPSWCWVNCQECISADTLVNYGTCMCLVGGIFYHCTNEDDEGS 218

Db 1 cegcgkckcgactaprltprlpsclacnrgclcsaesmevgtcmclvkglfthcsnddegds 60
 QY 219 CADHPCSCSRNCCARMSFNGALSVLPCLLCYLPATGCVKLAORGVRLRRPGRCRHT 278
 Db 61 ysdnpccsgshccserylcmgamslflpccllcypakgclklrrcydwlnpgrcckns 120
 QY 279 NSVICKAAS---GDAKYS 293
 Db 121 nlycklescparsgqkps 139

RESULT 13

ID AAW48794 standard; Protein: 117 AA.
 AC AAW48794;
 DT 11-SEP-1998 (first entry)

DE Homo sapiens sprouty 1 cysteine rich protein.

KW spry1; sprouty 1 protein; angiogenesis; treatment; cancer;
 KW lateral inhibitory signal; epithelial tube branching; inhibition;
 KW neovascularisation; diabetic retinopathy; retrolental fibroplasia;
 KW neovascular glaucoma; psoriasis; angiodiomas; immune; non-immune;
 KW inflammation; rheumatoid arthritis; capillary proliferation;
 KW atherosclerotic plaques; haemangiomas; Kaposi's Sarcoma;
 KW endometriosis; scar formation; wound healing.

OS Homo sapiens.

FX Key Location/Qualifiers
 FT Misc-difference 103
 FT /note="undefined amino acid"

W09820032-A1.

PD 14-MAY-1998.

PF 07-NOV-1997; 97WO-US20398.

PR 07-NOV-1996; 96US-0030232.

PA (STRD) UNIV LELAND STANFORD JUNIOR.

PI Hacothen N, Krasnow MA;

DR WPI; 1998-286862/25.

PT New isolated human sprouty polypeptide(s) - used to develop products
 PT for treating disorders involving angiogenic dysfunction, e.g.
 PT cancers, diabetic retinopathy, glaucoma, psoriasis or inflammation
 PS Claim 4: Page 52; 90pp: English.

CC The sequence is that of human sprouty 1 cysteine rich protein. It
 CC functions as a lateral inhibitory signal during epithelial tube
 CC branching. It can be used for inhibiting angiogenesis and
 CC neovascularisation in the treatment of e.g. cancers, diabetic
 CC retinopathy, retrolental fibroplasia, neovascular glaucoma, psoriasis,
 CC angiodiomas, immune and non-immune inflammation (including rheumatoid
 CC arthritis), capillary proliferation within atherosclerotic plaques,
 CC haemangiomas, Kaposi's Sarcoma, endometriosis and unwanted scar
 CC formation in wound healing. Products of the DNA and proteins
 CC can also be used for detection and diagnosis.

CC Sequence 117 AA;

Query Match 29.8%; Score 493; DB 19; Length 117;
 Best Local Similarity 66.7%; Pred. No. 6.9e-34;
 Matches 78; Conservative 21; Mismatches 18; Indels 0; Gaps 0;

QY 159 CEACGCKCKEACASPRTLPSGWCNOCCLCSAOTLVNVTGTCMLVGGIFHYCHTNEDEGS 218
 Db 1 cegcgkckcgactaprltprlpsclacnrgclcsaesmevgtcmclvkglfthcsnddegds 60
 QY 219 CADHPCSCSRNCCARMSFNGALSVLPCLLCYLPATGCVKLAORGVRLRRPGRCR 275
 Db 61 ysdnpccsgshccserylcmgamslflpccllcypakgclklrrcydwlnpgrcckns 117

RESULT 14

ID AAW48793 standard; Protein: 140 AA.
 AC AAW48793;
 DT 11-SEP-1998 (first entry)

DE Homo sapiens sprouty 1 protein.

KW spry1; sprouty 1 protein; angiogenesis; treatment; cancer;
 KW lateral inhibitory signal; epithelial tube branching; inhibition;
 KW neovascularisation; diabetic retinopathy; retrolental fibroplasia;
 KW neovascular glaucoma; psoriasis; angiodiomas; immune; non-immune;
 KW inflammation; rheumatoid arthritis; capillary proliferation;
 KW atherosclerotic plaques; haemangiomas; Kaposi's Sarcoma;
 KW endometriosis; scar formation; wound healing.

OS Homo sapiens.

FX W09820032-A1.

PD 14-MAY-1998.

PF 07-NOV-1997; 97WO-US20398.

PR 07-NOV-1996; 96US-0030232.

PA (STRD) UNIV LELAND STANFORD JUNIOR.

PI Hacothen N, Krasnow MA;

DR WPI; 1998-286862/25.

PT New isolated human sprouty polypeptide(s) - used to develop products
 PT for treating disorders involving angiogenic dysfunction, e.g.
 PT cancers, diabetic retinopathy, glaucoma, psoriasis or inflammation
 PS Disclosure; Page 50-51; 90pp: English.

CC The sequence is that of human sprouty 1 protein. The protein
 CC functions as a lateral inhibitory signal during epithelial tube
 CC branching. It can be used for inhibiting angiogenesis and
 CC neovascularisation in the treatment of e.g. cancers, diabetic
 CC retinopathy, retrolental fibroplasia, neovascular glaucoma, psoriasis,
 CC angiodiomas, immune and non-immune inflammation (including rheumatoid
 CC arthritis), capillary proliferation within atherosclerotic plaques,
 CC haemangiomas, Kaposi's Sarcoma, endometriosis and unwanted scar
 CC formation in wound healing. Products of the DNA and proteins
 CC can also be used for detection and diagnosis.

CC Sequence 140 AA;

Query Match 26.8%; Score 443.5; DB 19; Length 140;
 Best Local Similarity 55.1%; Pred. No. 1.2e-29;
 Matches 76; Conservative 21; Mismatches 22; Indels 19; Gaps 2;

QY 159 CEACGCKCKEACASPRTLPSGWCNOCCLCSAOTLVNVTGTCMLVGGIFHYCHTNEDEGS 218
 Db 1 cegcgkckcgactaprltprlpsclacnrgclcsaesmevgtcmclvkglfthcsnddegds 60

[illegible]

RESULT 15

ID	AAW46796 standard; Protein; 591 AA.
XX	
AC	AAW46796;
XX	
DT	11-SEP-1998 (first entry)
XX	
DE	Drosophila melanogaster sprouty protein.
XX	
KM	SPRY; sprouty protein.
XX	
OS	Drosophila melanogaster.
XX	
PN	W09820032-AI.
PD	
XX	14-MAY-1998.
XX	
PF	07-NOV-1997; 97WO-US20398.
XX	
PR	07-NOV-1996; 96US-0030232.
XX	
PA	(STRD) UNIV LELAND STANFORD JUNIOR.
XX	
PI	Hacohen N, Krasnow MA;
XX	
DR	WPI; 1998-286862/25.
DR	N-PSDB; AAV32402.
XX	
PT	New isolated human sprouty polypeptide(s) - used to develop products
XX	PT for treating disorders involving angiogenic dysfunction, e.g.
PT	cancers, diabetic retinopathy, glaucoma, psoriasis or inflammation
XX	
PS	Example 5; Fig 5C; 90pp; English.
CC	
CC	The sequence is that of Drosophila sprouty protein which was used
CC	in the isolation of human spry genes.
XX	
SO	Sequence 591 AA;

Query Match	23.9%	Score 396,	DB 19	Length 591,
Best Local Similarity	29.4%	Pred. No. 5.8e-25,		
Matches 95; Conservative	50;	Mismatches 120;	Indels 58;	Gaps 10;

```

QY 4 PIP0SARPTKPSVWVQORLPDLSRSMH-SRLQHLPLRLPDQKSHVENDYIDNPSLALT 62
Dp 190 presentlneuy-----drlqhaetqzhpqagqngtctthlllprqmgihlbg 243
QY 63 GKPKTRGGAERLAPTRARCDQDVTNNHITSRGRRSSVSSSSSTSSD-----QRLDHN 115
Dp 244 hqgdlqqgqgqgqgqgqghlqgqgqghatqatattqzswgsdhtdglhshlqnst 303
QY 116 APPRVADQ-ASPR-----AVRIGPKVVCORPLDK-----144
Dp 304 tkpraskqpralrfgmglgldglgldgldlqprlltkqpractqkethlaeelldprgagng 353
QY 145 GRAY----PRBLDKNHLTLCAGKCKKECASPTPLSCWVNOECLSAQTLVNRGTCM 2000
Dp 364 grlvmagdrslldp--lvrcpvcgctcegcqsrprldqvtwvsnktclsaesavldyascl 421
QY 201 CLVQGFHNCNEDQ-----DEGS---CADHNCSGRSKCARSWPMGLAVLPLCLSLPL 253
Dp 422 ccaakalfhncardlceddngprcvdnpdpsccrgykrctqagwlgalsflrplwtfwpr 461

```

QY 254 ATGCVKLAQRGYDRLRRPGCRCK 276
 ||:::| | |||:
 Db 482 mrgcmklcekygrfagrgcrg 504

Search completed: May 8, 2002, 10:46:15
Job time: 32 sec